PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07H 17/00, C12N 15/00, 5/00, C12Q 1/70, A61K 38/00, 35/14 — A1

(11) International Publication Number: WO 95/06055

(43) International Publication Date: 2 March 1995 (02.03.95)

(21) International Application Number: PCT/US94/09303

(22) International Filing Date: 19 August 1994 (19.08.94)

(30) Priority Data:
08/110,522 20 August 1993 (20.08.93) US
08/264,537 23 June 1994 (23.06.94) US

(60) Parent Application or Grant
(63) Related by Continuation
US 08/264,537 (CON)
Filed on 23 June 1994 (23.06.94)

(71) Applicant *(for all designated States except US)*: SMITHKLINE BEECHAM CORPORATION [US/US]; Corporate Intellectual Property, UW2220, 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406-0939 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: DILELLA, Anthony, G. [US/US]; 130 Hawthorne Lane, Phoenixville, PA 19460 (US). DEBOUCK, Christine, Marie [BE/US]; 667 Pugh Road, Wayne, PA 19087 (US).

(74) Agents: JERVIS, Herbert, H. et al.; SmithKline Beecham Corporation, Corporate Intellectual Property, UW2220, 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406-0939 (US).

(81) Designated States: AU, BB, BG, BR, BY, CA, CN, CZ, FI, HU, JP, KP, KR, KZ, LK, MG, MN, MW, NO, NZ, PL, RO, RU, SD, SI, SK, UA, US, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: HSV-2 UL26 GENE, CAPSID PROTEINS, IMMUNOASSAYS AND PROTEASE INHIBITORS

(57) Abstract

Essentially pure HSV-2 UL26 gene products and fragments thereof including mature HSV-2 protease and active fragments thereof are disclosed. Essentially pure HSV-2 UL26.5 gene products and fragments thereof including mature HSV-2 capsid protein and functional fragments are disclosed. Isolated nucleic acid molecules comprising all or part of the HSV-2 UL26 gene and/or the HSV-2 UL26.5 gene are disclosed. Expression vectors and host cells comprising such nucleic acid molecules are disclosed. Methods of identifying compounds that inhibit HSV-2 protease activity and methods of identifying compounds that inhibit HSV-2 virion assembly are disclosed. Synthetic HSV-2 substrates are disclosed. Antibodies that selectively bind to HSV-2 protease processed substrates but not unprocessed substrates or unprocessed substrates but not processed substrates are disclosed. Methods of and kits for distinguishing between HSV-1 DNA or protein and HSV-2 DNA or protein and reagents useful in such methods and kits are disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

HSV-2 UL26 GENE, CAPSID PROTEINS, IMMUNOASSAYS AND PROTEASE INHIBITORS

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of copending U. S. patent application serial number 08/110,522, filed August 20, 1993, the entire contents of which are incorporated herein by reference.

FIELD OF THE INVENTION

The present invention relates to HSV-2 UL26 and HSV-2 UL26.5 genes; to essentially pure HSV-2 UL26 and HSV-2 UL26.5 gene products; to compositions and methods of producing and using HSV-2 UL26 and HSV-2 UL26.5 DNA sequences and gene products.

BACKGROUND OF THE INVENTION

The herpes viruses consist of large icosahedral enveloped virions containing a linear double stranded genome. Currently, six human herpes viruses have been isolated and are known to be responsible for a variety of disease states from sub-clinical infections to fatal disease states in the immunocompromised. One human herpes virus, herpes simplex virus type 2, designated HSV-2, is usually acquired through sexual contact and gives rise to genital herpes. The frequency of recurrence of secondary genital herpes ranges between one and six times per year. It is estimated that genital HSV-2 infections occur in ten to sixty million individuals in the USA. Currently, there are no vaccines available to protect against HSV-2 infection.

Little is known regarding the genome composition of HSV-2. Nevertheless, HSV-2 presents a major public health problem. Individuals continue to become infected by the virus and no completely satisfactory anti-viral agents or vaccines are available. There is a need for a method of identifying anti-HSV-2 agents. There is a need for reagents useful in such methods. There is a need for a method of identifying compounds which modulate the activity of HSV-2 proteins and affect the ability of the virus to replicate and produce multiple infectious virions in an infected cell. There is a need for methods of and kits for distinguishing HSV-2 infections from other herpesvirus infections.

SUMMARY OF THE INVENTION

The present invention relates to essentially pure HSV-2 UL26 gene products and fragments thereof including HSV-2 protease precursor protein, mature HSV-2 protease and active fragments thereof, HSV capsid precursor protein and mature HSV-2 capsid protein.

The present invention relates to essentially pure HSV-2 UL26.5 gene products and fragments thereof including HSV-2 capsid precursor protein and mature HSV-2 capsid protein.

The present invention relates to isolated nucleic acid molecules comprising the HSV-2 UL26 gene or portions thereof including isolated nucleic acid molecules that encode mature HSV-2 protease and active fragments thereof and nucleic acid molesules that encode precursor or mature HSV-2 capsid protein, regulatory, e.g., promoter regions, or functional fragments thereof.

The present invention relates to expression vectors comprising the HSV-2 UL26 gene or portions thereof including nucleotide sequences that encode mature HSV-2 protease and active fragments thereof and nucleotide sequences that encode precursor or mature HSV-2 capsid protein or functional fragments thereof.

The present invention relates to host cells that contain expression vectors comprising the HSV-2 UL26 gene or portions thereof including nucleotide sequences that encode mature HSV-2 protease and active fragments thereof and nucleotide sequences that encode precursor or mature HSV-2 capsid protein or functional fragments thereof.

The present invention relates to isolated nucleic acid molecules comprising the HSV-2 UL26.5 gene or portions thereof including isolated nucleic acid molecules that encode mature HSV-2 capsid protein, regulatory, e.g., Promoter regions or fragments thereof and nucleotide sequences that encode precursor or mature HSV-2 capsid protein or functional fragments thereof.

The present invention relates to expression vectors comprising the HSV-2 UL26.5 gene or portions thereof including nucleotide sequences that encode mature HSV-2 capsid protein or fragments thereof and nucleotide sequences that encode precursor or mature HSV-2 capsid protein or functional fragments thereof.

The present invention relates to host cells that contain expression vectors comprising the HSV-2 UL26.5 gene or portions thereof including nucleotide sequences that encode mature HSV-2 capsid protein or fragments thereof and nucleotide sequences that encode precursor or mature HSV-2 capsid protein or functional fragments thereof.

The present invention relates to methods of identifying compounds that inhibit HSV-2 protease activity comprising contacting HSV-2 protease or active fragments thereof with an HSV-2 protease substrate in the presence of a test compound, detecting the level of proteolytic cleavage of the substrate and comparing that level to the level that occurs in the absence of the test compound.

The present invention relates to methods of identifying compounds that inhibit HSV-2 virion assembly by contacting HSV-2 capsid proteins in the presence of a test compound, detecting the level of capsid-capsid association and comparing that level to the level that occurs in the absence of the test compound.

The present invention relates to HSV-2 protease substrates produced by means of chemical synthesis or recombinantly produced and predicated on fragments or all of the UL26 gene product.

The present invention relates to antibodies that selectively bind to HSV-2 protease processed substrates but not unprocessed substrates or that selectively bind to unprocessed substrates but not to processed substrates.

The present invention relates to methods of distinguishing between HSV-1 DNA and HSV-2 DNA comprising PCR amplification of DNA using primers which will amplify HSV-1 DNA but not HSV-2 DNA and/or PCR amplification of DNA using primers which will amplify HSV-2 DNA but not HSV-1 DNA.

The present invention relates to PCR primers which will amplify HSV-1 DNA but not HSV-2 DNA and PCR primers which will amplify HSV-2 DNA but not HSV-1 DNA.

The present invention relates to kits for distinguishing between HSV-1 DNA and HSV-2 DNA comprising a container comprising PCR primers which will amplify HSV-1 DNA but not HSV-2 DNA and a positive control and size marker to determine if HSV-1 DNA has been amplified by the primers and/or a container comprising PCR primers which will amplify HSV-2 DNA but not HSV-1 DNA and a positive control and size marker to determine if HSV-2 DNA has been amplified by the primers.

The present invention relates to methods of distinguishing between HSV-1 protein and HSV-2 protein comprising an immunoassay using antibodies that selectively bind to HSV-1 protein but not HSV-2 protein and/or an immunoassay using antibodies that selectively bind to HSV-2 protein but not HSV-1 protein.

The present invention relates to antibodies which selectively bind to HSV-1 protein but not HSV-2 protein or antibodies which selectively bind to HSV-2 protein but not HSV-1 protein.

The present invention relates to kits for distinguishing between HSV-1 protein and HSV-2 protein. Said kit comprising a carrier being compartmented to receive a series of containers in close confinement which comprises a first container comprising antibodies which selectively bind to HSV-1 protein but not HSV-2 protein and a means to detect whether the antibodies are bound to HSV-1 protein and/or a second container comprising antibodies which selectively bind to HSV-2 protein but not HSV-1 protein and a means to detect whether the antibodies are bound to HSV-2 protein.

The present inventional relateds to the HSV-2 protease promoter and/or enhancer elements and their uses.

The present invention relates to the HSV-2 capsid protein promoter and/or enhancer elements and their uses.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 illustrates the HSV-2 UL26 gene. The symbol < > denotes the limits of the HSV-2 UL26 gene product. A putative termination codon is underlined.
The symbol [[ ]] denotes the limits of the HSV-2 UL26.5 gene product.
The symbol [ ] denotes the limits of two major proteolytic sites. The cissile bond is indicated by the *.
The symbol | | denotes the promoter region of the HSV-2 UL26.5 gene, a putative "TATA box" is underlined.
Figure 2 illustrates the expression of chloramphenicol acetyltransferase (CAT) when regulated in the HSV-2 UL26.5 promoter.

DETAILED DESCRIPTION OF THE INVENTION

As used herein, the term UL26 gene refers to a DNA molecule comprising a nucleotide sequence that encodes the HSV-2 protease and a form of the HSV-2 capsid protein. The UL26 gene is disclosed in SEQ ID NO:1. The coding region of the UL26 gene consists of nucleotides 534-2447 of SEQ ID NO:1. When expressed, the UL26 gene encodes a 638 amino acid active protease precursor disclosed in SEQ ID NO:1 and SEQ ID NO:2.

As used herein, the term "active protease precursor" refers to the unprocessed UL26 translation product. The active protease precursor is an active HSV-2 protease. When produced, the active protease precursor autocleaves at an internal protease cleavage site between amino acid residues 247 and 248. The amino terminal 247 amino acid portion retains protease activity.

As used herein, the term "mature protease" refers to the amino terminal 247 amino acid protein that is produced by autocleavage of the active protease precursor. The amino acid sequence of the mature protease is disclosed as amino acids 1-247 of SEQ ID NO:1 and SEQ ID NO:2.

As used herein, the term "HSV-2 protease" is meant to refer to, interchangeably, active protease precursor, mature protease or active fragments thereof.

As used herein, the term "UL26.5" gene refers to a DNA molecule comprising a nucleotide sequence that encodes the HSV-2 capsid protein. The UL26.5 gene is an internal sequence within the UL26 gene which is separately transcribed. The UL26.5 gene is disclosed in SEQ ID NO:1 and includes the coding region from nucleotide 1461-2447. When expressed, the UL26.5 gene encodes a 329 amino acid capsid precursor disclosed in SEQ ID NO:1 and SEQ ID NO:2 as amino acids 310-638.

As used herein, the term "capsid precursor" refers to the unprocessed UL26.5 translation product. While not wishing to be bound by any particular mechanistic theory regarding the function of the gene products of this invention, but based in part on the literature concerning HSV-1, it is believed that after it is produced, the capsid precursor is cleaved by the HSV-2 protease at an internal protease cleavage site between amino acid residues 613 and 614 of SEQ ID NO:1 and SEQ ID NO:2. The 304 amino acid portion is the capsid protein used in viral assembly and viral DNA packaging. It is the C-terminal processing of UL26.5 that enables packaging of viral DNA into mature capsids. Inhibition of this processing event results in the inability to package DNA into mature capsids.

As used herein, the term "mature capsid protein" refers to the 304 amino acid protein that is produced by cleavage of the capsid precursor by the HSV-2 protease. The amino acid sequence of the mature capsid protein is disclosed as amino acids 310-613 of SEQ ID NO:1 and SEQ ID NO:2.

As used herein, the term "HSV-2 capsid protein" is meant to refer to, interchangeably, capsid precursor and mature capsid protein.

As used herein the term "functional fragments" when used to modify a specific gene or gene product means a less than full length portion of the gene or gene product which retains substantially all of the biological function associated with the full length gene or gene product to which it relates. To determine whether a fragment of a particular gene or gene product is a functional fragment one merely

generates the fragments by well-known nucleolytic or proteolytic techniques and tests the thus generated fragments for the described biological function.

The present invention relates to essentially pure HSV-2 protease, to compositions and methods for producing and using HSV-2 protease, to nucleic acid molecules that encode HSV-2 protease and to methods for producing and using nucleic acid molecules that encode HSV-2 protease. The present invention relates to essentially pure HSV-2 capsid protein, to compositions and methods for producing and using HSV-2 capsid protein, to nucleic acid molecules that encode HSV-2 capsid protein, to methods for producing and using nucleic acid molecules that encode HSV-2 capsid protein. The present invention relates to substrates which are cleaved by HSV-2 protease, to methods of identifying compounds that inhibit HSV-2 protease activity, to methods of identifying compounds which inhibit HSV-2 capsid assembly, to methods of distinguishing between samples containing HSV-1 DNA and samples containing HSV-2 DNA, to methods of distinguishing between samples containing HSV-1 protein and samples containing HSV-2 protein, and to reagents, including oligonucleotides and antibodies, for performing such methods.

Some embodiments of the present invention provide methods for identifying compounds which inhibit or otherwise modulate the activity of HSV-2 protease. Thus, the present invention provides methods for identifying compounds useful as anti-HSV-2 agents since the activity of the HSV-2 protease is essential for the viral life cycle. According to the present invention, HSV-2 protease is contacted with an HSV-2 protease substrate (substrate) in the presence of a test compound to determine whether or not the test compound affects proteolytic activity. The effect of the test compound on the HSV-2 protease may be determined by comparing the proteolytic activity in the presence of the test compound to the proteolytic activity that would be observed in the absence of the compound.

Proteolytic activity refers to the ability of the HSV-2 protease to enzymatically process the substrate into products, i.e. cleave a single substrate peptide molecule into two or more peptide molecules (proteolytic products). In the viral life cycle, protease precursor is processed into mature protease and capsid precursor is processed into mature capsid by such proteolytic cleavage. This conversion is necessary for virion assembly and viral DNA packaging. The level of proteolytic activity may be determined by a variety of means well known by those having ordinary skill in the art. Essentially, a means is provided to distinguish unprocessed substrate from proteolytic product. Thus, after the substrate, HSV-2 protease and test compound are contacted, the level of proteolytic activity can be

observed by detecting the amount of unprocessed substrate remaining, the amount of unprocessed substrate depleted, or the amount of proteolytic product generated.

The present invention provides essentially pure HSV-2 protease which is useful in an assay to identify compounds which modulate HSV-2 protease activity. The present invention provides methods of producing essentially pure HSV-2 protease. The amino acid sequence of HSV-2 protease is disclosed in SEQ ID NO:1 and SEQ ID NO:2. As described above, the 638 amino acid active protease precursor is disclosed in SEQ ID NO:1 and SEQ ID NO:2. The active protease precursor is an active HSV-2 protease which is processed by autocleavage at an internal protease cleavage site between amino acid residues 247 and 248 to produce a 247 amino acid protein referred to as mature protease. Purified active protease precursor, mature protease and active fragments thereof may be produced by routine peptide synthesis methods or by using recombinant DNA technology using the information provided in SEQ ID NO:1. Using standard procedures and readily available starting materials, one having ordinary skill in the art can produce HSV-2 protease. Furthermore, using standard procedures and readily available starting materials, one having ordinary skill in the art can determine whether a fragment and/or derivative of the active protease precursor or mature protease is an active fragment.

Assays for determining whether or not a protein or peptide is capable of cleaving a specific substrate is disclosed herein. To determine if an HSV-2 protease fragment has proteolytic activity, one having ordinary skill in the art can perform protease activity assays as described herein without test compounds and using the fragment or derivative of the protease instead of the protease identical to SEQ ID NO:2. If the fragment or derivative cleaves the substrate, it is active, i.e. the fragment or derivative possesses proteolytic activity. Thus, one having ordinary skill in the art can routinely determine if a fragment or derivative of the protease is an active fragment or derivative.

The present invention relates to nucleotide sequences that encode HSV-2 protease and to nucleotide sequences that encode HSV-2 capsid protein. The UL26 gene including a nucleotide sequence which encodes HSV-2 protease and a precursor form of HSV-2 capsid protein is disclosed in SEQ ID NO:1. The UL26.5 gene including a nucleotide sequence which encodes HSV-2 capsid protein is also disclosed in SEQ ID NO:1. One having ordinary skill in the art can, using standard techniques and readily available starting materials, use the information disclosed herein including SEQ ID NO:1 to obtain or synthesize a nucleic acid molecule that

encodes HSV-2 protease or a nucleic acid molecule that encodes HSV-2 capsid protein. Further, using standard techniques, readily available starting materials and the information disclosed herein including SEQ ID NO:1, one having ordinary skill in the art can produce essentially pure HSV-2 protease including, active precursor protease, mature protease or active HSV-2 protease fragments. Likewise, using standard techniques, readily available starting materials and the information disclosed herein including SEQ ID NO:1, one having ordinary skill in the art can produce essentially pure HSV-2 capsid protein including capsid precursor, mature capsid, or HSV-2 capsid fragments capable of assembly functional fragments. One having ordinary skill in the art can, using standard techniques and readily available starting materials, use the information disclosed herein including SEQ ID NO:1 to obtain or synthesize a nucleic acid molecule that encodes HSV-2 protease or HSV-2 capsid protein using codons which provide optimum protein production in a given host cell used in an expression system.

Nucleic acid molecules encoding HSV-2 protease or HSV-2 capsid protein may be generated by those having ordinary skill in the art without undue experimentation using a variety of techniques. Using, for example, Polymerase Chain Reaction (PCR) methodology, primers may be designed and used to produce multiple copies of the nucleotide sequences that encode the HSV-2 protease or HSV-2 capsid protein. The entire nucleotide sequence encoding active protease precursor may be obtained routinely by amplifying the viral DNA. Similarly, the nucleotide sequence encoding mature protease may be obtained routinely by amplifying the viral DNA. Likewise, the nucleotide sequence encoding an active HSV-2 protease fragment may be obtained routinely by amplifying the viral DNA. In a similar manner, the entire nucleotide sequence encoding capsid precursor, mature capsid or functional fragments thereof may be obtained routinely by amplifying the viral DNA. Alternatively, using restriction enzymes, DNA encoding HSV-2 protease, including the active protease precursor, the mature protease, or active fragments thereof or HSV-2 capsid protein including capsid precursor, mature capsid or functional fragments thereof may be obtained from viral DNA cloned into vectors and identified by hybridization using probes designed from the disclosed nucleotide sequence. Moreover, nucleic acid molecules that encode the HSV-2 protease or the HSV-2 capsid protein may also be synthesized using techniques well known to those having ordinary skill in the art. Codons which encode HSV-2 protease or HSV-2 capsid protein may be selected to optimize protein production in a host cell selected for recombinant production of the HSV-2 protease or HSV-2

capsid protein. The HSV-2 genome is highly rich in G+C nucleotides. This is particularly true for the UL26 gene which encodes HSV-2 protease. Such high G+C character poses a problem in overexpressing genes in *E. coli* because of codon usage and an increased chance of frame-shift mutations. In an effort to improve expression of UL26 in *E. coli*, the UL26 gene and fragments thereof were changed to provide codons preferred in *E. coli* yet maintaining the authentic amino acid sequence of the protease. The reference for preferred codon usage is: Wada *et al.*, (1992) "Codon Usage Tabulated from the GenBank Genetic Sequence Data", *Nucleic Acid Research*, Vol. 20 Supplement, pages 2111-2118, which is incorporated herein by reference. Optimization of codon usage is well known and can be employed to design nucleic acid molecules according to the present invention which can be expressed at an improved level of efficiency in a selected host.

One having ordinary skill in the art can, using well known techniques, insert such DNA molecules into vectors such as commercially available expression vectors for use in well known expression systems. For example, commercially available plasmids such as pSE420 (Invitrogen, San Diego, CA) or pET-16(b) (Novagen, Madison W.I.) may be used for production of HSV-2 protease in *E. coli*. The commercially available plasmid pYES2 (Invitrogen, San Diego, CA) may, for example, be used for production in *S. cerevisiae* strains of yeast. The commercially available MAXBAC™ complete baculovirus expression system (Invitrogen, San Diego, CA) may, for example, be used for production in insect cells. The commercially available plasmid pcDNA I (Invitrogen, San Diego, CA) may, for example, be used for production in mammalian cells such as Chinese Hamster Ovary cells. One having ordinary skill in the art can use these commercial expression vectors and systems or others to produce the HSV-2 protease or HSV-2 capsid protein using routine techniques and readily available starting materials. (*See e.g.*, Sambrook et al., *Molecular Cloning a Laboratory Manual*, Second Ed. Cold Spring Harbor Press (1989) which is incorporated herein by reference.) Thus, the desired proteins can be prepared in both prokaryotic and eukaryotic systems, resulting in a spectrum of processed forms of the protein.

The particulars for the construction of expression systems suitable for desired hosts are known to those in the art. Briefly, for recombinant production of the protein, the DNA encoding the polypeptide is suitably ligated into the expression vector of choice. The DNA is operably linked to all regulatory elements which are necessary for expression of the DNA in the selected host. One having ordinary skill

in the art can, using well known techniques, prepare expression vectors for recombinant production of the polypeptide.

The expression vector including the DNA that encodes the HSV-2 protease or HSV-2 capsid protein is used to transform or transfect the compatible host which is then cultured and maintained under conditions wherein expression of the foreign DNA takes place. The protein of the present invention thus produced is recovered from the culture, either by lysing the cells or from the culture medium as appropriate and known to those in the art. One having ordinary skill in the art can, using well known techniques, isolate the protein that is produced using such expression systems.

According to one embodiment of the invention, protein may be produced and purified as follows. A DNA molecule that comprises a nucleotide sequence that encodes the HSV-2 protease or the HSV-2 capsid protein is produced which includes a nucleotide sequence that encodes multiple histidine residues at a terminal portion of the protein. This DNA molecule is incorporated into an expression vector which is introduced into suitable host cells. The DNA is expressed and the protein, including the terminal histidine residues, which are referred to herein as the histidine tag or His-tag, is produced. The cells are collected and maintained on ice in phosphate buffered saline at pH 8.5. The cells are then lysed by sonication. The sonicated cellular material is centrifuged at 30,000 x g. The supernatant is then filtered through a .2 micron filter. The filtered supernatant is incubated with a metal chelating resin (e.g., a nitrilo triacetic acid nickel resin is one of many such resins useful for such a purpose) for 2 hours at room temperature, after which time the resin is separated from unbound material by centrifugation. The resin is then packed into a column and washed with 50 mM imidazole to eliminate non specifically bound proteins. The His-tagged protease is then eluted from the Ni column with 150 mM imidazole buffer. The eluate from the column is further purified by column chromatography using Pharmacia Superdex 75 sizing column in phosphate buffered saline.

The DNA molecule may be engineered to include a specific cleavage site between the histidine tag and authentic HSV-2 protease to enable removal of the histidine tag from the expressed protein. Removal of the histidine tag may be accomplished as follows: The $(asparte)_4$ lysine sequence can be engineered to follow the histidine tag and precede the HSV-2 sequence when the histine tag is placed at the amino-terminus of the HSV-2 protease. Enterokinase specifically

cleaves after the (aspartate)$_4$lysine sequence thereby generating authentic HSV-2 protease.

In addition to producing these proteins by recombinant techniques, automated peptide synthesizers may also be employed to produce the HSV-2 protease or the HSV-2 capsid protein. Such techniques are well known to those having ordinary skill in the art.

The present invention provides essentially pure substrates for HSV-2 protease cleavage activity including synthetic substrates. An HSV-2 protease substrate is a peptide which can be cleaved at least into two separate peptides by HSV-2 protease mediated proteolysis. In some embodiments, the size differential between cleaved and uncleaved substrates may be used to detect whether or not the protease is active. In some embodiments, the substrates of the present invention are labelled so that they may be detected. In some embodiments, the substrates are fixed to a solid phase. In some embodiments of the invention, either the substrate or a proteolytic product has a biologically or chemical activity not present in the other which can be used to distinguish one from the other. Examples of biological activities include enzyme activity and the ability to bind with specific antibodies.

Two amino acid sequences are contained in UL26 that have been identified as natural cleavage sites. The first is LQAS (SEQ ID NO:3) wherein the HSV-2 protease cleaves the peptide between the A and the S. The second is VNAS (SEQ ID NO:4) wherein the HSV-2 protease cleaves the peptide between the A and the S. Natural or synthetic substrates may be produced which contain either of these two cleavage sites. Accordingly, a substrate according to the present invention have either the formula

$$R_1 - \text{SEQ ID NO:3} - R_2$$

or the formula

$$R_1 - \text{SEQ ID NO:4} - R_2$$

wherein R_1 and R_2 are, independently, hydrogen or one or more amino acids. In some embodiments, the substrate is the UL26 gene product which contains two protease cleavage sites: one comprising SEQ ID NO:3 and one comprising SEQ ID NO:4. In some embodiments, the substrate is the UL26.5 gene product which contains a protease cleavage sites comprising SEQ ID NO:4. In some embodiments, R_1 is preferably 1-20 amino acids, more preferably 1-10, and most preferably 3, 4, 5, 6, 7, 8 or 9 amino acids. In some embodiments, R_2 is preferably 1-20 amino acids, more preferably 1-10, and most preferably 3, 4, 5, 6, 7, 8 or 9 amino acids.

One having ordinary skill in the art can readily design substrates according to the above formula. The following peptides have been designed as substrates.

1. Peptides including the internal cleavage site SEQ ID NO:3 (LQA*S):

AHTYLQA*SEKFK	SEQ ID NO:5
AGIAGHTYLQA*SEKFK	SEQ ID NO:6
GIAGHTYLQA*SEKFK	SEQ ID NO:7
IAGHTYLQA*SEKFK	SEQ ID NO:8
GHTYLQA*SEKFK	SEQ ID NO:9
HTYLQA*SEKFKM	SEQ ID NO:10
HTYLQA*SEKFKMW	SEQ ID NO:11
HTYLQA*SEKFKMWG	SEQ ID NO:12
HTYLQA*SEKFKMWGA	SEQ ID NO:13
HTYLQA*SEKFKMWGAE	SEQ ID NO:14

2. Peptides including the terminal cleavage site SEQ ID NO:4 (VNA*S):

ALVNA*SSAAHVDVD	SEQ ID NO:15

The asterisk (*) indicates the scissile bond where cleavage by HSV-2 protease occurs.

The substrates may be obtained from proteolytic cleavage of the UL26 or UL26.5 protein product. They may be produced recombinantly by expression of UL26 or UL26.5 gene or fragment thereof containing the cleavage site or may be made by means of synthetic organic chemical means using standard peptide synthetic procedures well known in the art such as Merrifield synthesis.

One having ordinary skill in the art can readily design assays using the HSV-2 protease and substrate to identify compounds that modulate HSV-2 protease activity. As used herein, the term "test assay" refers to assays that include a mixture of HSV-2 protease, substrate and test compound; and the term "control assay" refers to assays that include a mixture of HSV-2 protease and substrate without test compound. To determine whether or not a test compound modulates HSV-2 protease activity, the level of HSV-2 protease activity in a test assay may be compared to the level of HSV-2 protease activity in a control assay.

In some embodiments of the present invention, the size differential between cleaved and uncleaved substrate is used to determine whether or not substrates are cleaved when contacted with HSV-2 protease in the presence of a test compound. In some embodiments, an HPLC assay is performed. Sample containing protease is

incubated with a substrate, for example HTYLQASEKFKMWGAE (SEQ ID NO:14), for 4 hrs at 37° C in phosphate buffered saline after which the reaction is terminated with trifluoroacetic acid. The reaction is then run on an HPLC column, showing activity manifested by the peptide cleavage products.

In some embodiments of the present invention, immunoassays are used to detect whether or not substrates are cleaved when contacted with HSV-2 protease in the presence of a test compound. In some embodiments, antibodies are provided which specifically bind to uncleaved substrate but not HSV-2 protease cleavage products. Such antibodies are referred to herein as "substrate-specific antibodies". In some embodiments, antibodies are provided which specifically bind to HSV-2 protease cleavage products but not uncleaved substrate. Such antibodies are referred to herein as "product-specific antibodies". Antibodies which react to either a product or a substrate but not both (i.e. substrate-specific antibodies and product-specific antibodies collectively) are referred to herein as "non-crossreactive antibodies". In some embodiments, antibodies are fixed to a solid phase. In some embodiments, antibodies are labelled.

For example, a mixture containing HSV-2 protease, substrate and test compound is maintained under appropriate conditions and for a sufficient amount of time to allow the proteolytic reaction to occur unless the test compound affects the reaction. The mixture can be added to a container which has non-crossreactive antibodies attached to the inner surface. If the non-crossreactive are substrate-specific antibodies, any uncleaved substrate remaining in the mixture will bind to the antibodies. If the substrate is labelled, the contained may be rinsed and the amount of label present may be detected. The level of HSV-2 protease activity is determined accordingly. If the non-crossreactive are product-specific antibodies, any HSV-2 protease products in the mixture will bind to the antibodies. If the substrate is labelled at a portion which is liberated as the product, the contained may be rinsed and the amount of label present may be detected. The level of HSV-2 protease activity is determined accordingly.

ICP35 antibodies (Catalog No.: 13-118-100; Rivers Park, 9108 Gulford Rd. Columbia, Maryland) may be used to detect cleaved substrate. Such antibodies are product specific and only bind to capsid protein after it has been proteolytically processed by the HSV-2 protease.

Alternatively, instead of using labelled substrates, the exemplified immunoassays may be modified as sandwich assays in which antibodies specific for the bound antigen complex are detected. Such antibodies are referred to herein as

complex-specific antibodies. The container is again rinsed and sufficient time is allowed for the binding of the complex specific antibody to any complex present. The level of complex specific antibody is detected and indicative of the level of HSV-2 protease activity.

In some embodiments of immunoassays, unlabelled substrate is used in the reaction mixture. After the reaction mixture is added to a container comprising a non-crossreactive antibody and maintained for a sufficient time for the non-crossreactive antibody to bind to either substrate or product, either labelled substrate or labelled product, respectively, is added and will bind to any non-crossreactive antibody not bound with substrate or product from the reaction mixture. Detecting the amount of labelled substrate or labelled product indicates the level of proteolytic cleavage.

In some embodiments, the substrate is labeled and the label is released when the substrate is converted to proteolytic products. Detecting the release of the label, which indicates the HSV-2 protease activity, may be accomplished by a variety of well known means. In some embodiments, labelled substrate is fixed to a solid phase. Upon cleavage by HSV-2 protease, the label attached to the portion of the substrate that becomes an unattached product, is released. Comparing the level of label present before and after the reaction mixture indicates how much label is released and thus the level of HSV-2 protease activity. Alternatively, detecting the amount of label freed from the solid phase indicates the level of HSV-2 protease activity.

In another embodiment, methods of detecting HSV-2 protease activity include fluorescence liberation assays in which substrate contains fluorescent label adjacent to the scissile bond. At such a location, the label is not detectable in uncleaved substrate. However, when the substrate is cleaved by HSV-2 protease at the cleavage site, the fluorescent group becomes exposed and the fluorescence becomes detectable. Thus, the level of proteolytic activity may be measured by measuring detectable fluorescence after contacting the substrate with HSV-2 protease in the presence of a test compound.

In another embodiment, methods of detecting HSV-2 protease activity include scintillation proximity assays in which radiolabelled substrate is conjugated to solid beads which, when in close proximity to the radiolabel, are excited and become detectable by scintillation. When the substrate is cleaved, the radiolabel is no longer in close proximity to the beads and the beads are not excited and not detectable by scintillation. Thus, the level of proteolytic activity may be measured

by measuring the excitation of the beads by scintillation after contacting the conjugated substrate with HSV-2 protease in the presence of a test compound.

In addition to these embodiments, one having ordinary skill in the art can apply well known techniques to devise other methods of identifying compounds that modulate HSV-2 protease activity using various means of detecting the HSV-2 protease cleavage or the lack thereof.

The present invention relates to kits for identifying compounds that modulate HSV-2 protease activity. Such kits include separate containers which comprise HSV-2 protease, substrate, and optionally, antibodies or other reagents for detecting HSV-2 protease activity or distinguishing between uncleaved substrate and products. The substrate or antibodies may be fixed to the inner surface of a container. The substrate or antibodies may be labelled.

Some embodiments of the present invention also provide methods of identifying compounds which inhibit or otherwise modulate HSV-2 capsid assembly using a multimerization assay. The present invention provides methods of identifying compounds useful as anti-HSV-2 agents since capsid assembly is essential for viral replication and infectivity. According to the present invention, chimeric genes are provided which comprise either a sequence including the HSV-2 UL26.5 gene or a portion thereof which encodes an HSV-2 capsid protein linked to a sequence encoding the yeast *GAL4* DNA-binding protein or a sequence including the HSV-2 UL26.5 gene or a portion thereof which encodes an HSV-2 capsid protein linked to a sequence encoding the yeast *GAL4* activation protein. While it is preferred that the portion of the chimeric gene that encodes the HSV-2 capsid protein encodes the mature capsid, the capsid precursor protein may also be usefully employed. Chimeric genes are inserted into *Saccharomyces cerevisiae* plasmids and the plasmids are introduced in *S. cerevisiae* which contains an integrated *GAL4*-responsive *lacZ* indicator gene. When the chimeric genes on the plasmids are expressed, fusion proteins are produced. The portions of the fusion proteins comprising the HSV-2 capsid protein will, under selected condition bind to each other and thereby bring together the DNA-binding domain and activation domain of *GAL4*. When the two *GAL4* domains which are in close proximity interact with the *GAL4*-responsive *lacZ* indicator gene, the indicator gene is expressed and, under the proper conditions a detectable blue color is observed. If the fusion proteins are prevented from binding, the two *GAL4* domains will not be present in proximity to each other and the indicator gene will not be activated. Thus, no blue color will be present to observe.

Thus, this yeast system provides a rapid and specific assay for the interaction of HSV-2 capsid proteins that occur during virion assembly. In the presence of compounds which interrupt or inhibit HSV-2 capsid protein interaction, the *GAL4* domains in the fusion proteins produced by expression of the chimeric genes will not associate and thereby will not activate the *lacZ* gene in the yeast system. Accordingly, compounds may be identified by the absence of activation of the *lacZ* gene in transformed yeast which inhibit HSV-2 capsid assembly and therefore possess anti-viral properties.

Some embodiments of the present invention provides methods of distinguishing between samples containing HSV-1 DNA and samples containing HSV-2 DNA or samples containing HSV-1 proteins and samples containing HSV-2 proteins. Accordingly, the present invention provides a method of diagnosing whether an individual is infected with HSV-1 and/or HSV-2. Methods are disclosed for identifying whether an individual is infected with HSV-1 and/or HSV-2 wherein HSV-1 infection can be distinguished from HSV-2 infection.

According to some embodiments of the invention, PCR technology is used to distinguish between samples containing HSV-1 DNA and samples containing HSV-2 DNA. Such methods provide a means for distinguishing between HSV-1 and HSV-2 infections and allow for the diagnosis of the type of HSV infection an individual has. Specific primers are designed that will provide for amplification of HSV-1 DNA but not HSV-2 DNA and/or HSV-2 DNA but not HSV-1 DNA. Accordingly, by performing amplification techniques using such primers with biological samples taken from individuals such as cell, serum or tissue samples, especially samples taken at sites where blisters or other manifestations of viral shedding are observed, one can determine whether or not the DNA in the sample is derived from HSV-1 or HSV-2 and therefore whether the individual from which the sample was taken is infected with HSV-1 or HSV-2.

The nucleotide sequence of the UL26 gene including the nucleotide sequence which encodes the HSV-2 protease and the HSV-2 capsid protein is disclosed in SEQ ID NO:1. The nucleotide sequence encoding HSV-1 protease and HSV-1 capsid protein are disclosed in SEQ ID NO:16. A set of PCR primers were designed which amplify HSV-2 sequences but not HSV-1 sequences. Thus, detection of amplified DNA indicates that HSV-2 is present. Similarly, a set of PCR primers were designed which amplify HSV-1 sequences but not HSV-2 sequences. Thus, detection of amplified DNA indicates that HSV-1 is present. It is preferred that both sets of primers are provided and used in separate amplification

protocols with material from the same sample in order to provide an additional control. Other optional controls include positive controls which contain DNA sequences that will be amplified and/or negative controls that cannot be amplified by the primers. Amplified DNA may be detected by running the material on an electrophoresis gel after the amplification protocol is complete. DNA molecules of the expected length of an amplification product may be provided as size markers.

Present invention also relates to kits for distinguishing whether a sample contains DNA from HSV-1 or HSV-2. The kits of the present invention are useful to diagnose whether an individual is infected with HSV-1 and/or HSV-2. The kits contain containers which comprise primers that will amplify HSV-1 DNA but not HSV-2 DNA or containers that will amplify HSV-2 DNA but not HSV-1 DNA. Kits may optionally contain both sets of primers in separate containers for running separate amplification procedures using different portions of the same sample. Kits may optionally contain positive and/or negative controls in separate containers. Kits may optionally contain DNA molecules in a separate container which can serve as a size marker. The DNA molecule may be of the expected length of a DNA molecule amplified using the primers.

According to some embodiments of the invention, immunoassays are used to distinguish between samples containing HSV-1 protein and samples containing HSV-2 protein. The immunoassays are used to distinguish between HSV-1 and HSV-2 infections and diagnose the type of HSV infection an individual has. Such immunoassays are based upon differences between UL26 gene products of HSV-1 and HSV-2 or between UL26.5 gene products of HSV-1 and HSV-2. Immunoassays may be based upon differences in proteases and/or capsid proteins of HSV-1 and HSV-2. Specific antibodies are provided which selectively bind to epitopes on HSV-1 antigens not present on HSV-2 antigens or which selectively bind to epitopes on HSV-2 antigens not present on HSV-1 antigens. For example, specific antibodies are provided which selectively bind to HSV-1 protease but not HSV-2 protease or which selectively bind to HSV-2 protease but not HSV-1 protease. Likewise, specific antibodies are provided which selectively bind to HSV-1 capsid but not HSV-2 capsid or which selectively bind to HSV-2 capsid but not HSV-1 capsid.

Accordingly, by performing antibody binding assays, using specific antibodies with biological samples taken from individuals such as cell, serum or tissue samples, especially samples taken from sites where blisters or other manifestations of viral shedding are observed, one can determine whether or not the

HSV-1-specific antibodies or the HSV-2-specific antibodies bind to proteins in the sample and therefore whether the individual from which the sample was taken is infected with HSV-1 and/or HSV-2. The amino acid sequence of HSV-2 active protease precursor spans amino acids 1-638 in SEQ ID NO:1 and SEQ ID NO:2. The amino acid sequence of HSV-2 mature protease spans amino acids 1-247 of SEQ ID NO:1 and SEQ ID NO:2. The amino acid sequence of HSV-2 capsid precursor spans amino acids 310-638 in SEQ ID NO:1 and SEQ ID NO:2. The amino acid sequence of HSV-2 mature capsid spans amino acids 310-613 of SEQ ID NO:1 and SEQ ID NO:2. The amino acid sequence of HSV-1 protease and capsid are disclosed in SEQ ID NO:17. The amino acid sequence of HSV-1 active protease precursor spans amino acids 1-635 in SEQ ID NO:17. The amino acid sequence of HSV-1 mature protease spans amino acids 1-247 of SEQ ID NO:17. The amino acid sequence of HSV-1 capsid precursor spans amino acids 307-635 in SEQ ID NO:17. The amino acid sequence of HSV-1 mature capsid spans amino acids 307-610 of SEQ ID NO:17.

Antibodies which specifically bind to HSV-2 protease but not HSV-1 protease may be produced by those having ordinary skill in the art using routine methods and widely available starting materials. Likewise, antibodies which specifically bind to HSV-2 capsid but not HSV-1 capsid may be produced by those having ordinary skill in the art using routine methods and widely available starting materials. Either of these HSV-2 specific antibodies are used to detect HSV-2 in an immunoassay which can distinguish HSV-1 from HSV-2. Similarly, antibodies which specifically bind to HSV-1 protease but not HSV-2 protease may be produced by those having ordinary skill in the art using routine methods and widely available starting materials. Likewise, antibodies which specifically bind to HSV-1 capsid but not HSV-2 capsid may be produced by those having ordinary skill in the art using routine methods and widely available starting materials. These HSV-1 specific antibodies are used to detect HSV-1 in an immunoassay which can distinguish HSV-1 from HSV-2. It is preferred that both immunoassays be performed using material from the same sample in order to provide an additional control. Other optional controls include positive controls which include peptides which will bind to the antibody used in the immunoassay and/or negative controls which include peptides which will not bind to the antibody used in the immunoassay. Antibodies may be labelled. Alternatively, an antibody that specifically binds to the HSV specific antibodies may be used. One having ordinary

skill in the art can readily produce immunoassays including all necessary reagents using the information provided herein.

HSV-1 protease antibody produced by Serotech as Antibody 45KD and commercially available from Bioproducts for Science Inc. as catalog number MCA406 (P.O. Box 29176, Indianapolis, IN) can be used in immunoassays to distinguish HSV-2 from HSV-1. The Serotech antibody binds to HSV-1 precursor or mature capsid protein but not HSV-2 precursor or mature capsid protein. Accordingly, an immunoassay using the Serotech antibodies may be performed to determine if a sample contains HSV-1 or HSV-2 and thus if the individual from whom the sample was taken is infected with HSV-1 or HSV-2.

Present invention also relates to kits for diagnosing whether an individual is infected with HSV-1 or HSV-2. The kits of the present invention may comprise a container comprising antibodies which bind to HSV-1 protease but not HSV-2 protease and/or a container comprising antibodies which bind to HSV-2 protease but not HSV-1 protease. It is preferred that the kit comprises both types of antibodies in separate containers. Antibodies used in the kits may be labelled. The kits contain all other reagents and materials for performing an immunoassay with the antibodies. Kits may optionally contain positive and/or negative controls in separate containers. Kits may optionally contain means to detect the antibody including, for example a second antibody which specifically binds to the anti-HSV protease antibody. The kits of the present invention may comprise a container comprising antibodies which bind to HSV-1 capsid but not HSV-2 capsid and/or a container comprising antibodies which bind to HSV-2 capsid but not HSV-1 capsid. It is preferred that the kit comprises both types of antibodies in separate containers. Antibodies used in the kits may be labelled. The kits contain all other reagents and materials for performing an immunoassay with the antibodies. Kits may optionally contain positive and/or negative controls in separate containers. Kits may optionally contain means to detect the antibody including, for example a second antibody which specifically binds to the anti-HSV capsid antibody. Kits may comprise the Serotech antibody.

Another aspect of the present invention relates to the HSV-2 protease promoter and/or enhancer elements and their uses. The HSV-2 protease promoter may be synthesized or isolated and linked to coding sequences which encode proteins other than HSV-2 protease. Accordingly, the present invention relates to recombinant DNA molecules which comprise at least a portion of the nucleotide sequence between nucleotides 1-534 of SEQ ID NO:1 operably linked to a

nucleotide sequence that encodes a protein other than HSV-2 protease. The present invention relates to cells which comprise DNA molecules which comprise at least a portion of the nucleotide sequence between 1 and 534 of SEQ ID NO:1 operably linked to a nucleotide sequence that encodes a protein other than HSV-2 protease.

Another aspect of the invention applies to bacteriophage lambda clones which harbor HSV-2 UL26 gene (SEQ. I.D. No.:1) and sequences upstream and downstream of the gene. Accordingly, the linked sequences can be used to screen for UL26 promoter regulatory and/or enhancer regions.

Another aspect of the present invention relates to the HSV-2 capsid protein promoter and its uses. The HSV-2 capsid protein promoter is located upstream of nucleotide 1461 of SEQ ID NO:1. It may be synthesized or isolated and linked to coding sequences which encode proteins other than HSV-2 capsid protein. Accordingly, the present invention relates to recombinant DNA molecules which comprise at least a portion of the nucleotide sequence upstream of nucleotide 1461 of SEQ ID NO:1 operably linked to a nucleotide sequence that encodes a protein other than HSV-2 capsid protein. The present invention relates to cells which comprise DNA molecules which comprise at least a portion of the nucleotide sequence upstream of nucleotide 1461 of SEQ ID NO:1 operably linked to a nucleotide sequence that encodes a protein other than HSV-2 capsid protein. Nucleotides 1191 to 1461 (SEQ ID NO:1), for example, were linked to the chloramphenicol acetyl transferase gene and shown to possess significant promoter activity when transfected into VERO cells

EXAMPLES

Example 1

A proteolytic activity essential to the virion maturation of herpes viruses has been characterized for HSV-2. The HSV-2 protease, also referred to as HSV-2 UL26, has a molecular weight (Apparent) of about 67,028 Da and a pI = 6.94. The HSV-2 protease can be employed using molecular and biochemical technology in *in vitro* assays identify inhibitors of this activity by rational design and screening and to test these inhibitors for antiviral activity in infected cells.

The HSV-2 UL26 gene was cloned as an *Nco*I-*Eco*RI fragment (1938 base pairs) which contained the start codon, the entire open reading frame, the stop codon, and 22 base pairs of 3'-untranslated sequence. Full-length HSV-2 UL26 was expressed in *E. coli* using the pOTS vector system in which the gene is inserted downstream of the strong and tightly regulated P_L promoter from bacteriophage

lambda of the p0TS-207 vector. Tight regulation of the promoter is essential when expressing genes that are likely to be toxic to the cells, such as proteases. The 27 KD protease domain corresponding to one of the autoproteolytic products derived from the HSV-2 UL26 primary translation product was produced in *E. coli* using the tightly regulated expression vector pET-16(b) (Novagen, Madison W.I.) which contains the T7 promoter.

Each construct was designed to include six histidine codons and the $(aspartate)_4$lysine codons preceding the HSV-2 UL26 start codon so that the expressed protein will contain a cleavable histidine tag at the N-terminus for purification of the protein on Nickel columns. Other chelating columns may be used. The His-tagged protein is eluted from the column by addition of imidazole Alternatively, it can be eluted by other means such as pH change. Columns and technical protocols useful to purify protein may be obtained from commercially available sources such as Qiagen.

For the P_L promoter vectors, the recombinant constructs are then introduced into *E. coli* AR120 (nalidixic acid inducible strain) and *E. coli* AR58 (heat inducible strain) for expression and processing/ purification studies. For the T7 promoter vectors, the recombinant constructs are introduced into *E. coli* BL21, an IPTG inducible strain. The proteins can be readily purified by chromatography on nickel chelate column.

The p27 protease fragment is active as shown by its ability to remove the last 25 amino acids from a construct comprising most of the UL26.5 coding region.

Example 2

The p27 protease gene was synthesized to contain codons characteristic of highly expressed E. coli genes, yet maintaining the amino acid sequence of p27 protease. The synthesized gene was placed downstream of the tightly regulated T7 promoter in the expression vector pET-16(b). Following IPTG (1mM) induction the 27 k Da protein domain was highly expressed in E. coli.

Example 3

The above HSV-2 UL26 gene (*Nco*I-*Eco*RI fragment) and the p27 protease is cloned into the insect cell expression vector pVL1392. The recombinant construct is then introduced into insect cells derived from *Spodoptera frugiperda.*

High titer viral stocks are then prepared for protease activity analysis and subsequent scale up for protein production.

Example 4

Oligonucleotide PCR primers were designed to the DNA region of HSV-1 UL26 gene and HSV-2 UL26 gene that shared the least amount of identity to ensure the specificity of the assay. Such a region can easily be viewed by computer analysis comparing the two DNA sequences disclosed in SEQ ID NO:1 and SEQ ID NO:16, respectively. The region of least identity between the two homologs lies within the UL26.5 domain, i.e. the portion of the gene that encodes the capsid. The following provides the sequences of the primers used and the locations of the primers are given based on the nucleotide numbers given in the nucleotide sequence comparison provided in the enclosed computer analysis. As shown below it is helpful to design a system to generate HSV-1 and HSV-2 specific products of different sizes to improve the analysis.

5'-PCR primers (sense-strand sequence):

SEQ ID NO:18 HSV-1: 5'-CCGGTGCCCAATCGTCCGT-3' (#864-882)

SEQ ID NO:19 HSV-2: 5'-GTCCGTGCGCGTCAAGTCG-3' (#1397-1416)

3'-PCR primers (antisense-strand sequence):

SEQ ID NO:20 HSV-1: 5'-TTCCGGCTCCCCCACCTGA-3' (#1560-1542)

SEQ ID NO:21 HSV-2: 5'-ATTCGGATCCTGGAGGCGA-3' (#2470-2452)

Expected PCR product sizes using these sets of primers:

HSV-1: 696 base pairs

HSV-2: 1073 base pairs.

Separate PCR amplification protocols are performed on samples suspected of containing either HSV-1 or HSV-2 DNA using SEQ ID NO:18 and SEQ ID NO:20 in the HSV-1 assay or SEQ ID NO:19 and SEQ ID NO:21 in the HSV-2 assay. If a DNA fragment of 696 base pairs is generated in the HSV-1 assay, the presence of HSV-1 DNA in the sample is indicated. To detect the presence of a 696 base pair fragment, the amplification product is migrated through an electrophoresis matrix. A size marker of DNA of about 696 base pairs is run through the same matrix simultaneously. If a DNA fragment of 1073 base pairs is generated in the HSV-2 assay, the presence of HSV-2 DNA in the sample is indicated. To detect the

presence of a 1073 base pair fragment, the amplification product is migrated through an electrophoresis matrix. A size marker of DNA of about 1073 base pairs is run through the same matrix simultaneously.

A kit is provided which comprises a container comprising SEQ ID NO:18 and SEQ ID NO:20 in the HSV-1 assay. A kit is provided which comprises a container comprising SEQ ID NO:19 and SEQ ID NO:21 in the HSV-2 assay. A kit is provided which comprises both a container comprising SEQ ID NO:18 and SEQ ID NO:20 in the HSV-1 assay and a container comprising SEQ ID NO:19 and SEQ ID NO:21 in the HSV-2 assay. Size marker DNA may optionally be provided. In some kits, a size marker of 696 base pairs is provided. In some kits, a size marker of 1073 base pairs is provided. In some kits, a size marker of 696 base pairs and a size marker of 1073 base pairs are provided.

Example 5

A region of the putative HSV-2 UL26.5 promoter contained in the HSV-2 UL26 gene was cloned to test for promoter activity. The 256 base pair region that was analyzed spanned nucleotides #1191 to # 1447 of SEQ ID NO:1. The DNA fragment was cloned by the polymerase chain reaction using the sense strand primer (5'-AACATGAGCTGCGTGACC-3') spanning nucleotide # 1191 to # 1209 of SEQ ID NO: 1 and the antisense strand primer (5'-AAAGAAGAAGAAGAAGAC-3') spanning nucleotides #1447 to # 1429 of SEQ ID NO: 1 Promoter activity is tested by cloning the 256 base pair PCR fragment upstream of the chloramphenicol acetyltransferase (CAT) reporter gene in the commercially available vector pCAT Basic (Promega). The resulting construct can then be introduced into a suitable mammalian cell line, e.g., Vero cells, to test for promoter activity by analyzing the levels of CAT activity. The cell line is devoid of endogenous CAT activity; hence, after introducing the promoter construct into such a cell line, the levels of CAT activity is a direct measure of HSV-2 UL26.5 promoter activity.

Vero cells were grown in DMEM+10% FCS containing Gentamicine (10ug/ml). 15 micograms of the HSV-2 UL26.5/pCAT construct was electroporated into 5 million Vero cells using standard protocols. 48 hrs after electroporation cells were harvested in 100 microliters of 0.25 M Tris buffer pH 8.0. Cells were lysed by repeated freeze-thaw, spun down at 15,000 rpm and the supernatants were transferred to fresh tubes. Total protein concentration was determined using Bio-Rad Protein Assay Dye Reagent Kit (Cat. # 500-0006). 5 Microliters of D-

Threo[dichloroacetyl-1-^{14}C] Chloramphenicol (Amersham, 56 mCi/mmol) and 5 microliters of n-Butyral CoA (5 mg/ml) was added to an aliquot of cell extract supernatant in a 100 microliter final volume to assay for CAT activity by ethyl acetate extraction followed by thin layer chromatography.

In addition to the above construct, the 256 base pair HSV-2 UL 26.5 fragment was also cloned upstream of the CAT reporter gene in the pCAT Enhancer vector, which contains an SV40 enhancer element. This construct was also tested for CAT activity in Vero cells by the same methods described above.

The control vector pCAT control (contains the SV40 promoter and enhancer) was used as a comparison of HSV-2 UL26.5 promoter strength.

Figure 2 summarizes the results of four experiments. Column 1 is a negative control and represents CAT expression in the absense of promoter and enhancer transcriptional control elements. Column 2, a positive control, employs SV40 promoter and SV40 enhancer elements to drive CAT gene expression. Column 3 represents CAT gene expression driven by UL26.5 promoter alone and Column 4 represents CAT gene expression when the UL26.5 promoter is used in combination with the SV40 enhancer element.

Having established a basal UL26.5 expression level (Column 3), additional fragments of the gene sequence within figure 1 can be used to identify the UL26.5 enhancer elements merely by isolating fragments of convenient length upstream from nucleotide 1191 back to nucleotide 1, introducing the fragments into the basal expression construct oriented operatively with respect to the promoter region and testing their ability to enhance CAT expression over the basal level.

The promoter described here are useful for regulating the expression of heterologous genes when operably linked thereto.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: DiLella, Anthony G.
Debouck, Christine

(ii) TITLE OF INVENTION: Novel Gene

(iii) NUMBER OF SEQUENCES: 21

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: SmithKline Beecham Corporation
Corporate Patents - US UW2220
(B) STREET: P.O. Box 1539
(C) CITY: King of Prussia
(D) STATE: Pennsylvania
(E) COUNTRY: USA
(F) ZIP: 19406-0939

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25, mmd

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Jervis, Herbert H.
(B) REGISTRATION NUMBER: 31,171
(C) REFERENCE/DOCKET NUMBER: P50188

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 215-270-5019
(B) TELEFAX: 215-270-5090

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2472 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 534..2447

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1461..2447

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GTCGACGAGG CGCGTGGTGG ATATGTCGTC GGGCGCCCGC CAGGCGGCGC TCGTGCGCCT      60

CACCGCGCTG GAGCTCATCA ACCGCACCCG CACAAACACC ACCCCTGTGG GGGAGATTAT     120

TAACGCCCAC GATGCCTTGG GGATACAATA CGAACAGGGC CTGGGGCTGC TCGCCCAGCA     180

GGCACGCATC GGCTTGGCGT CGAACGCCAA GCGATTCGCC ACGTTCAACG TGGGCAGCGA     240

CTACGACCTG TTGTACTTTT TGTGTCTCGG GTTCATTCCC CAGTACCTGT CCGTGGCCTA     300

GGGAAGGGTG GGGGTGGTGG TGGTGGGGTG TTTTTCTGTT GTTGTTGTTT CTGGTCCGCC     360

TGGTCACAAA AGGCACGGCG CCCCGAAACG CGGGCTTTAG TCCCGGCCCG GACGTCGGCG     420

GACACACAAC AACGGCGGGC CCCGTGGGTG GGTAAGTTGG TTCGGGGGCA TCGCTGTATT     480

CCCTTGCCCG CTTCCACCCC CCCTTCCCGT TTGGTTTGTT TGTGCGGGTG CCC ATG        536
                                                            Met
                                                             1

GCG TCG GCG GAA ATG CGC GAG CGG TTG GAG GCG CCT CTG CCC GAC CGG        584
Ala Ser Ala Glu Met Arg Glu Arg Leu Glu Ala Pro Leu Pro Asp Arg
          5                  10                  15

GCG GTG CCC ATC TAC GTG GCC GGG TTT TTG GCC CTG TAC GAC AGC GGG        632
Ala Val Pro Ile Tyr Val Ala Gly Phe Leu Ala Leu Tyr Asp Ser Gly
     20                  25                  30

GAC CCG GGC GAG CTG GCC CTG GAC CCA GAC ACG GTG CGT GCG GCC CTG        680
Asp Pro Gly Glu Leu Ala Leu Asp Pro Asp Thr Val Arg Ala Ala Leu
35                  40                  45

CCT CCG GAG AAC CCC CTG CCG ATC AAC GTA GAC CAC CGC GCT CGG TGC        728
Pro Pro Glu Asn Pro Leu Pro Ile Asn Val Asp His Arg Ala Arg Cys
50                  55                  60                  65

GAG GTG GGC CGG GTG CTC GCC GTG GTC AAC GAC CCT CGG GGG CCG TTT        776
Glu Val Gly Arg Val Leu Ala Val Val Asn Asp Pro Arg Gly Pro Phe
                70                  75                  80

TTT GTG GGG CTG ATC GCG TGC GTG CAG CTG GAG CGC GTC CTC GAG ACG        824
Phe Val Gly Leu Ile Ala Cys Val Gln Leu Glu Arg Val Leu Glu Thr
          85                  90                  95

GCC GCC AGC GCC GCT ATT TTT GAG CGC CGC GGA CCC GCG CTC TCC CGG        872
Ala Ala Ser Ala Ala Ile Phe Glu Arg Arg Gly Pro Ala Leu Ser Arg
     100                 105                 110

GAG GAG CGT CTG CTG TAC CTG ATC ACC AAC TAC CTG CCA TCG GTC TCG        920
Glu Glu Arg Leu Leu Tyr Leu Ile Thr Asn Tyr Leu Pro Ser Val Ser
115                 120                 125

CTG TCC ACA AAA CGC CGG GGG GAC GAG GTT CCG CCC GAC CGC ACC CTG        968
Leu Ser Thr Lys Arg Arg Gly Asp Glu Val Pro Pro Asp Arg Thr Leu
130                 135                 140                 145

TTT GCG CAC GTG GCC CTG TGC GCC ATC GGG CGG CGC CTT GGA ACC ATC       1016
Phe Ala His Val Ala Leu Cys Ala Ile Gly Arg Arg Leu Gly Thr Ile
                150                 155                 160
```

```
GTC ACC TAC GAC ACC AGC CTA GAC GCG GCC ATC GCT CCG TTT CGC CAC      1064
Val Thr Tyr Asp Thr Ser Leu Asp Ala Ala Ile Ala Pro Phe Arg His
            165                 170                 175

CTG GAC CCG GCG ACG CGC GAG GGG GTG CGA CGC GAG GCC GCC GAG GCC      1112
Leu Asp Pro Ala Thr Arg Glu Gly Val Arg Arg Glu Ala Ala Glu Ala
        180                 185                 190

GAG CTC GCG CTG GCC GGG CGC ACC TGG GCC CCC GGC GTG GAG GCG CTC      1160
Glu Leu Ala Leu Ala Gly Arg Thr Trp Ala Pro Gly Val Glu Ala Leu
    195                 200                 205

ACA CAC ACG CTG CTC TCC ACC GCC GTC AAC AAC ATG ATG CTG CGT GAC      1208
Thr His Thr Leu Leu Ser Thr Ala Val Asn Asn Met Met Leu Arg Asp
210                 215                 220                 225

CGC TGG AGC CTC GTG GCC GAG CGG CGG CGG CAG GCC GGG ATC GCC GGA      1256
Arg Trp Ser Leu Val Ala Glu Arg Arg Arg Gln Ala Gly Ile Ala Gly
                230                 235                 240

CAC ACG TAC CTT CAG GCG AGC GAA AAA TTT AAA ATA TGG GGG GCG GAG      1304
His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys Ile Trp Gly Ala Glu
            245                 250                 255

TCT GCC CCT GCG CCG GAG CGT GGG TAT AAA ACC GGC GCC CCG GGT GCC      1352
Ser Ala Pro Ala Pro Glu Arg Gly Tyr Lys Thr Gly Ala Pro Gly Ala
        260                 265                 270

ATG GAC ACA TCC CCC GCC GCG AGC GTT CCC GCG CCG CAG GTC GCC GTC      1400
Met Asp Thr Ser Pro Ala Ala Ser Val Pro Ala Pro Gln Val Ala Val
    275                 280                 285

CGT GCG CGT CAA GTC GCG TCG TCG TCG TCT TCT TCT TCT TCT TTT CCG      1448
Arg Ala Arg Gln Val Ala Ser Ser Ser Ser Ser Ser Ser Ser Phe Pro
290                 295                 300                 305

GCA CCG GCC GAT ATG AAC CCC GTT TCG GCA TCG GGC GCC CCG GCC CCT      1496
Ala Pro Ala Asp Met Asn Pro Val Ser Ala Ser Gly Ala Pro Ala Pro
                310                 315                 320

CCG CCG CCC GGC GAC GGG AGT TAT TTG TGG ATC CCC GCC TCT CAT TAC      1544
Pro Pro Pro Gly Asp Gly Ser Tyr Leu Trp Ile Pro Ala Ser His Tyr
            325                 330                 335

AAT CAG CTC GTC ACC GGG CAA TCC GCG CCC CGC CAC CCG CCG CTG ACC      1592
Asn Gln Leu Val Thr Gly Gln Ser Ala Pro Arg His Pro Pro Leu Thr
        340                 345                 350

GCG TGC GGC CTG CCG GCC GCG GGG ACG GTG GCC TAC GGA CAC CCC GGC      1640
Ala Cys Gly Leu Pro Ala Ala Gly Thr Val Ala Tyr Gly His Pro Gly
    355                 360                 365

GCC GGC CCG TCC CCG CAC TAC CCG CCT CCT CCC GCC CAC CCG TAC CCG      1688
Ala Gly Pro Ser Pro His Tyr Pro Pro Pro Pro Ala His Pro Tyr Pro
370                 375                 380                 385

GGT ATG CTG TTC GCG GGC CCC AGT CCC CTG GAG GCC CAG ATC GCC GCG      1736
Gly Met Leu Phe Ala Gly Pro Ser Pro Leu Glu Ala Gln Ile Ala Ala
                390                 395                 400

CTG GTG GGG GCC ATC GCC GCC GAC CGC CAG GCG GGT GGG CTT CCG GCG      1784
Leu Val Gly Ala Ile Ala Ala Asp Arg Gln Ala Gly Gly Leu Pro Ala
            405                 410                 415
```

```
GCC GCC GGA GAC CAC GGG ATC CGG GGG TCG GCG AAG CGC CGC CGA CAC      1832
Ala Ala Gly Asp His Gly Ile Arg Gly Ser Ala Lys Arg Arg Arg His
        420                 425                 430

GAG GTG GAG CAG CCG GAG TAC GAC TGC GGC CGT GAC GAG CCG GAC CGG      1880
Glu Val Glu Gln Pro Glu Tyr Asp Cys Gly Arg Asp Glu Pro Asp Arg
    435                 440                 445

GAC TTC CCG TAT TAC CCG GGC GAG GCC CGC CCC GAG CCG CGC CCG GTC      1928
Asp Phe Pro Tyr Tyr Pro Gly Glu Ala Arg Pro Glu Pro Arg Pro Val
450                 455                 460                 465

GAC TCC CGG CGC GCC GCG CGC CAG GCT TCC GGG CCC CAC GAA ACC ATC      1976
Asp Ser Arg Arg Ala Ala Arg Gln Ala Ser Gly Pro His Glu Thr Ile
                470                 475                 480

ACG GCG CTG GTG GGG GCG GTG ACG TCC CTG CAG CAG GAA CTG GCG CAC      2024
Thr Ala Leu Val Gly Ala Val Thr Ser Leu Gln Gln Glu Leu Ala His
            485                 490                 495

ATG CGC GCG CGT ACC CAC GCC CCC TAC GGG CCG TAT CCG CCG GTG GGG      2072
Met Arg Ala Arg Thr His Ala Pro Tyr Gly Pro Tyr Pro Pro Val Gly
        500                 505                 510

CCC TAC CAC CAC CCC CAC GCA GAC ACG GAG ACC CCC GCC CAA CCA CCC      2120
Pro Tyr His His Pro His Ala Asp Thr Glu Thr Pro Ala Gln Pro Pro
    515                 520                 525

CGC TAC CCC GCC GAG GCC GTC TAT CTG CCG CCG CCG CAC ATC GCC CCC      2168
Arg Tyr Pro Ala Glu Ala Val Tyr Leu Pro Pro Pro His Ile Ala Pro
530                 535                 540                 545

CCG GGG CCT CCT CTA TCC GGG GCG GTC CCC CCA CCC TCG TAT CCC CCA      2216
Pro Gly Pro Pro Leu Ser Gly Ala Val Pro Pro Pro Ser Tyr Pro Pro
                550                 555                 560

GTT GCG GTT ACC CCC GGT CCC GCT CCC CCG CTA CAT CAG CCC TCC CCC      2264
Val Ala Val Thr Pro Gly Pro Ala Pro Pro Leu His Gln Pro Ser Pro
            565                 570                 575

GCA CAC GCC CAC CCC CCT CCG CCG CCG CCG GGA CCC ACG CCT CCC CCC      2312
Ala His Ala His Pro Pro Pro Pro Pro Pro Gly Pro Thr Pro Pro Pro
        580                 585                 590

GCC GCG AGC TTA CCC CAA CCC GAG GCG CCC GGC GCG GAG GCC GGC GCC      2360
Ala Ala Ser Leu Pro Gln Pro Glu Ala Pro Gly Ala Glu Ala Gly Ala
    595                 600                 605

TTA GTT AAC GCC AGC AGC GCG GCC CAC GTG AAC GTG GAC ACG GCC CGG      2408
Leu Val Asn Ala Ser Ser Ala Ala His Val Asn Val Asp Thr Ala Arg
610                 615                 620                 625

GCC GCC GAT CTG TTT GTG TCA CAG ATG ATG GGG TCC CGC TAACTCGCCT       2457
Ala Ala Asp Leu Phe Val Ser Gln Met Met Gly Ser Arg
                630                 635

CCAGGATCCG AATTC                                                    2472
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 638 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Ala Ser Ala Glu Met Arg Glu Arg Leu Glu Ala Pro Leu Pro Asp
 1               5                  10                  15

Arg Ala Val Pro Ile Tyr Val Ala Gly Phe Leu Ala Leu Tyr Asp Ser
            20                  25                  30

Gly Asp Pro Gly Glu Leu Ala Leu Asp Pro Asp Thr Val Arg Ala Ala
        35                  40                  45

Leu Pro Pro Glu Asn Pro Leu Pro Ile Asn Val Asp His Arg Ala Arg
    50                  55                  60

Cys Glu Val Gly Arg Val Leu Ala Val Val Asn Asp Pro Arg Gly Pro
65                  70                  75                  80

Phe Phe Val Gly Leu Ile Ala Cys Val Gln Leu Glu Arg Val Leu Glu
                85                  90                  95

Thr Ala Ala Ser Ala Ala Ile Phe Glu Arg Arg Gly Pro Ala Leu Ser
            100                 105                 110

Arg Glu Glu Arg Leu Leu Tyr Leu Ile Thr Asn Tyr Leu Pro Ser Val
        115                 120                 125

Ser Leu Ser Thr Lys Arg Arg Gly Asp Glu Val Pro Pro Asp Arg Thr
    130                 135                 140

Leu Phe Ala His Val Ala Leu Cys Ala Ile Gly Arg Arg Leu Gly Thr
145                 150                 155                 160

Ile Val Thr Tyr Asp Thr Ser Leu Asp Ala Ala Ile Ala Pro Phe Arg
                165                 170                 175

His Leu Asp Pro Ala Thr Arg Glu Gly Val Arg Arg Glu Ala Ala Glu
            180                 185                 190

Ala Glu Leu Ala Leu Ala Gly Arg Thr Trp Ala Pro Gly Val Glu Ala
        195                 200                 205

Leu Thr His Thr Leu Leu Ser Thr Ala Val Asn Asn Met Met Leu Arg
    210                 215                 220

Asp Arg Trp Ser Leu Val Ala Glu Arg Arg Arg Gln Ala Gly Ile Ala
225                 230                 235                 240

Gly His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys Ile Trp Gly Ala
                245                 250                 255

Glu Ser Ala Pro Ala Pro Glu Arg Gly Tyr Lys Thr Gly Ala Pro Gly
            260                 265                 270

Ala Met Asp Thr Ser Pro Ala Ala Ser Val Pro Ala Pro Gln Val Ala
        275                 280                 285
```

```
Val Arg Ala Arg Gln Val Ala Ser Ser Ser Ser Ser Ser Ser Ser Phe
290                 295                 300

Pro Ala Pro Ala Asp Met Asn Pro Val Ser Ala Ser Gly Ala Pro Ala
305                 310                 315                 320

Pro Pro Pro Pro Gly Asp Gly Ser Tyr Leu Trp Ile Pro Ala Ser His
                325                 330                 335

Tyr Asn Gln Leu Val Thr Gly Gln Ser Ala Pro Arg His Pro Pro Leu
            340                 345                 350

Thr Ala Cys Gly Leu Pro Ala Ala Gly Thr Val Ala Tyr Gly His Pro
        355                 360                 365

Gly Ala Gly Pro Ser Pro His Tyr Pro Pro Pro Pro Ala His Pro Tyr
370                 375                 380

Pro Gly Met Leu Phe Ala Gly Pro Ser Pro Leu Glu Ala Gln Ile Ala
385                 390                 395                 400

Ala Leu Val Gly Ala Ile Ala Ala Asp Arg Gln Ala Gly Gly Leu Pro
                405                 410                 415

Ala Ala Ala Gly Asp His Gly Ile Arg Gly Ser Ala Lys Arg Arg Arg
            420                 425                 430

His Glu Val Glu Gln Pro Glu Tyr Asp Cys Gly Arg Asp Glu Pro Asp
        435                 440                 445

Arg Asp Phe Pro Tyr Tyr Pro Gly Glu Ala Arg Pro Glu Pro Arg Pro
450                 455                 460

Val Asp Ser Arg Arg Ala Ala Arg Gln Ala Ser Gly Pro His Glu Thr
465                 470                 475                 480

Ile Thr Ala Leu Val Gly Ala Val Thr Ser Leu Gln Gln Glu Leu Ala
                485                 490                 495

His Met Arg Ala Arg Thr His Ala Pro Tyr Gly Pro Tyr Pro Pro Val
            500                 505                 510

Gly Pro Tyr His His Pro His Ala Asp Thr Glu Thr Pro Ala Gln Pro
        515                 520                 525

Pro Arg Tyr Pro Ala Glu Ala Val Tyr Leu Pro Pro Pro His Ile Ala
530                 535                 540

Pro Pro Gly Pro Pro Leu Ser Gly Ala Val Pro Pro Pro Ser Tyr Pro
545                 550                 555                 560

Pro Val Ala Val Thr Pro Gly Pro Ala Pro Pro Leu His Gln Pro Ser
                565                 570                 575

Pro Ala His Ala His Pro Pro Pro Pro Pro Pro Pro Gly Pro Thr Pro Pro
            580                 585                 590

Pro Ala Ala Ser Leu Pro Gln Pro Glu Ala Pro Gly Ala Glu Ala Gly
        595                 600                 605

Ala Leu Val Asn Ala Ser Ser Ala Ala His Val Asn Val Asp Thr Ala
610                 615                 620
```

Arg Ala Ala Asp Leu Phe Val Ser Gln Met Met Gly Ser Arg
625 630 635

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Leu Gln Ala Ser
1

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Val Asn Ala Ser
1

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Ala His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys
1 5 10

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Ala Gly Ile Ala Gly His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Gly Ile Ala Gly His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Ile Ala Gly His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys
1 5 10

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Gly His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys
1 5 10

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys Met
1 5 10

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys Met Trp
1 5 10

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys Met Trp Gly
1 5 10

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys Met Trp Gly Ala
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys Met Trp Gly Ala Glu
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Ala Leu Val Asn Ala Ser Ser Ala Ala His Val Asp Val Asp
1 5 10

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1908 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..1908

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 919..1908

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```
ATG GCA GCC GAT GCC CCG GGA GAC CGG ATG GAG GAG CCC CTG CCC GAC   48
Met Ala Ala Asp Ala Pro Gly Asp Arg Met Glu Glu Pro Leu Pro Asp
 1               5                  10                  15

AGG GCC GTG CCC ATT TAC GTG GCT GGG TTT TTG GCC CTG TAT GAC AGC   96
Arg Ala Val Pro Ile Tyr Val Ala Gly Phe Leu Ala Leu Tyr Asp Ser
            20                  25                  30

GGG GAC TCG GGC GAG TTG GCA TTG GAT CCG GAT ACG GTG CGG GCG GCC  144
Gly Asp Ser Gly Glu Leu Ala Leu Asp Pro Asp Thr Val Arg Ala Ala
        35                  40                  45

CTG CCT CCG GAT AAC CCA CTC CCG ATT AAC GTG GAC CAC CGC GCT GGC  192
Leu Pro Pro Asp Asn Pro Leu Pro Ile Asn Val Asp His Arg Ala Gly
    50                  55                  60

TGC GAG GTG GGG CGG GTG CTG GCC GTG GTC GAC GAC CCC CGC GGG CCG  240
Cys Glu Val Gly Arg Val Leu Ala Val Val Asp Asp Pro Arg Gly Pro
65                  70                  75                  80

TTT TTT GTG GGG CTG ATC GCC TGC GTG CAG CTG GAG CGC GTC CTC GAG  288
Phe Phe Val Gly Leu Ile Ala Cys Val Gln Leu Glu Arg Val Leu Glu
                85                  90                  95

ACG GCC GCC AGC GCT GCG ATT TTC GAG CGC CGC GGG CCG CCG CTC TCC  336
Thr Ala Ala Ser Ala Ala Ile Phe Glu Arg Arg Gly Pro Pro Leu Ser
            100                 105                 110

CGG GAG GAG CGC CTG TTG TAC CTG ATC ACC AAC TAC CTG CCC TCG GTC  384
Arg Glu Glu Arg Leu Leu Tyr Leu Ile Thr Asn Tyr Leu Pro Ser Val
        115                 120                 125

TCC CTG GCC ACA AAA CGC CTG GGG GGC GAG GCG CAC CCC GAT CGC ACG  432
Ser Leu Ala Thr Lys Arg Leu Gly Gly Glu Ala His Pro Asp Arg Thr
    130                 135                 140

CTG TTC GCG CAC GTC GCG CTG TGC GCG ATC GGG CGG CGC CTC GGC ACT  480
Leu Phe Ala His Val Ala Leu Cys Ala Ile Gly Arg Arg Leu Gly Thr
145                 150                 155                 160

ATC GTC ACC TAC GAC ACC GGT CTC GAC GCC GCC ATC GCG CCC TTT CGC  528
Ile Val Thr Tyr Asp Thr Gly Leu Asp Ala Ala Ile Ala Pro Phe Arg
                165                 170                 175

CAC CTG TCG CCG GCG TCT CGC GAG GGG GCG CGG CGA CTG GCC GCC GAG  576
His Leu Ser Pro Ala Ser Arg Glu Gly Ala Arg Arg Leu Ala Ala Glu
            180                 185                 190

GCC GAG CTC GCG CTG TCC GGG CGC ACC TGG GCG CCC GGC GTG GAG GCG  624
Ala Glu Leu Ala Leu Ser Gly Arg Thr Trp Ala Pro Gly Val Glu Ala
        195                 200                 205

CTG ACC CAC ACG CTG CTT TCC ACC GCC GTT AAC AAC ATG ATG CTG CGG  672
Leu Thr His Thr Leu Leu Ser Thr Ala Val Asn Asn Met Met Leu Arg
    210                 215                 220

GAC CGC TGG AGC CTG GTG GCC GAG CGG CGG CGG CAG GCC GGG ATC GCC  720
Asp Arg Trp Ser Leu Val Ala Glu Arg Arg Arg Gln Ala Gly Ile Ala
225                 230                 235                 240
```

```
GGA CAC ACC TAC CTC CAG GCG AGC GAA AAA TTC AAA ATG TGG GGG GCG      768
Gly His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys Met Trp Gly Ala
                245                 250                 255

GAG CCT GTT TCC GCG CCG GCG CGC GGG TAT AAG AAC GGG GCC CCG GAG      816
Glu Pro Val Ser Ala Pro Ala Arg Gly Tyr Lys Asn Gly Ala Pro Glu
            260                 265                 270

TCC ACG GAC ATA CCG CCC GGC TCG ATC GCT GCC GCG CCG CAG GGT GAC      864
Ser Thr Asp Ile Pro Pro Gly Ser Ile Ala Ala Ala Pro Gln Gly Asp
        275                 280                 285

CGG TGC CCA ATC GTC CGT CAG CGC GGG GTC GCC TTG TCC CCG GTA CTG      912
Arg Cys Pro Ile Val Arg Gln Arg Gly Val Ala Leu Ser Pro Val Leu
    290                 295                 300

CCC CCC ATG AAC CCC GTT CCG ACA TCG GGC ACC CCG GCC CCC GCG CCG      960
Pro Pro Met Asn Pro Val Pro Thr Ser Gly Thr Pro Ala Pro Ala Pro
305                 310                 315                 320

CCC GGC GAC GGG AGC TAC CTG TGG ATC CCG GCC TCC CAT TAC AAC CAG     1008
Pro Gly Asp Gly Ser Tyr Leu Trp Ile Pro Ala Ser His Tyr Asn Gln
                325                 330                 335

CTC GTC GCC GGC CAT GCC GCG CCC CAA CCC CAG CCG CAT TCC GCG TTT     1056
Leu Val Ala Gly His Ala Ala Pro Gln Pro Gln Pro His Ser Ala Phe
            340                 345                 350

GGT TTC CCG GCT GCG GCG GGG TCC GTG GCC TAT GGG CCT CAC GGT GCG     1104
Gly Phe Pro Ala Ala Ala Gly Ser Val Ala Tyr Gly Pro His Gly Ala
        355                 360                 365

GGT CTT TCC CAG CAT TAC CCT CCC CAC GTC GCC CAT CAG TAT CCC GGG     1152
Gly Leu Ser Gln His Tyr Pro Pro His Val Ala His Gln Tyr Pro Gly
    370                 375                 380

GTG CTG TTC TCG GGA CCC AGC CCA CTC GAG GCG CAG ATA GCC GCG TTG     1200
Val Leu Phe Ser Gly Pro Ser Pro Leu Glu Ala Gln Ile Ala Ala Leu
385                 390                 395                 400

GTG GGG GCC ATA GCC GCG GAC CGC CAG GCG GGC GGT CAG CCG GCC GCG     1248
Val Gly Ala Ile Ala Ala Asp Arg Gln Ala Gly Gly Gln Pro Ala Ala
                405                 410                 415

GGA GAC CCT GGG GTC CGG GGG TCG GGA AAG CGT CGC CGG TAC GAG GCG     1296
Gly Asp Pro Gly Val Arg Gly Ser Gly Lys Arg Arg Arg Tyr Glu Ala
            420                 425                 430

GGG CCG TCG GAG TCC TAC TGC GAC CAG GAC GAA CCG GAC GCG GAC TAC     1344
Gly Pro Ser Glu Ser Tyr Cys Asp Gln Asp Glu Pro Asp Ala Asp Tyr
        435                 440                 445

CCG TAC TAC CCC GGG GAG GCT CGA GGC GCG CCG CGC GGG GTC GAC TCC     1392
Pro Tyr Tyr Pro Gly Glu Ala Arg Gly Ala Pro Arg Gly Val Asp Ser
    450                 455                 460

CGG CGC GCG GCC CGC CAT TCT CCC GGG ACC AAC GAG ACC ATC ACG GCG     1440
Arg Arg Ala Ala Arg His Ser Pro Gly Thr Asn Glu Thr Ile Thr Ala
465                 470                 475                 480

CTG ATG GGG GCG GTG ACG TCT CTG CAG CAG GAA CTG GCG CAC ATG CGG     1488
Leu Met Gly Ala Val Thr Ser Leu Gln Gln Glu Leu Ala His Met Arg
                485                 490                 495
```

```
GCT CGG ACC AGC GCC CCC TAT GGA ATG TAC ACG CCG GTG GCG CAC TAT      1536
Ala Arg Thr Ser Ala Pro Tyr Gly Met Tyr Thr Pro Val Ala His Tyr
            500                 505                 510

CGC CCT CAG GTG GGG GAG CCG GAA CCA ACA ACG ACC CAC CCG GCC CTT      1584
Arg Pro Gln Val Gly Glu Pro Glu Pro Thr Thr Thr His Pro Ala Leu
        515                 520                 525

TGT CCC CCG GAG GCC GTG TAT CGC CCC CCA CCA CAC AGC GCC CCC TAC      1632
Cys Pro Pro Glu Ala Val Tyr Arg Pro Pro Pro His Ser Ala Pro Tyr
    530                 535                 540

GGT CCT CCC CAG GGT CCG GCG TCC CAT GCC CCC ACT CCC CCG TAT GCC      1680
Gly Pro Pro Gln Gly Pro Ala Ser His Ala Pro Thr Pro Pro Tyr Ala
545                 550                 555                 560

CCA GCT GCC TGC CCG CCA GGC CCG CCA CCG CCC CCA TGT CCT TCC ACC      1728
Pro Ala Ala Cys Pro Pro Gly Pro Pro Pro Pro Pro Cys Pro Ser Thr
                565                 570                 575

CAG ACG CGC GCC CCT CTA CCG ACG GAG CCC GCG TTC CCC CCC GCC GCC      1776
Gln Thr Arg Ala Pro Leu Pro Thr Glu Pro Ala Phe Pro Pro Ala Ala
            580                 585                 590

ACC GGA TCC CAA CCG GAG GCA TCC AAC GCG GAG GCC GGG GCC CTT GTC      1824
Thr Gly Ser Gln Pro Glu Ala Ser Asn Ala Glu Ala Gly Ala Leu Val
        595                 600                 605

AAC GCC AGC AGC GCA GCA CAC GTG GAC GTT GAC ACG GCC CGC GCC GCC      1872
Asn Ala Ser Ser Ala Ala His Val Asp Val Asp Thr Ala Arg Ala Ala
    610                 615                 620

GAT TTG TTC GTC TCT CAG ATG ATG GGG GCC CGC TGA                      1908
Asp Leu Phe Val Ser Gln Met Met Gly Ala Arg
625                 630                 635
```

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 635 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```
Met Ala Ala Asp Ala Pro Gly Asp Arg Met Glu Glu Pro Leu Pro Asp
 1               5                  10                  15

Arg Ala Val Pro Ile Tyr Val Ala Gly Phe Leu Ala Leu Tyr Asp Ser
            20                  25                  30

Gly Asp Ser Gly Glu Leu Ala Leu Asp Pro Asp Thr Val Arg Ala Ala
        35                  40                  45

Leu Pro Pro Asp Asn Pro Leu Pro Ile Asn Val Asp His Arg Ala Gly
    50                  55                  60

Cys Glu Val Gly Arg Val Leu Ala Val Val Asp Asp Pro Arg Gly Pro
65                  70                  75                  80
```

```
Phe Phe Val Gly Leu Ile Ala Cys Val Gln Leu Glu Arg Val Leu Glu
                85                  90                  95

Thr Ala Ala Ser Ala Ala Ile Phe Glu Arg Arg Gly Pro Pro Leu Ser
            100                 105                 110

Arg Glu Glu Arg Leu Leu Tyr Leu Ile Thr Asn Tyr Leu Pro Ser Val
        115                 120                 125

Ser Leu Ala Thr Lys Arg Leu Gly Gly Glu Ala His Pro Asp Arg Thr
    130                 135                 140

Leu Phe Ala His Val Ala Leu Cys Ala Ile Gly Arg Arg Leu Gly Thr
145                 150                 155                 160

Ile Val Thr Tyr Asp Thr Gly Leu Asp Ala Ala Ile Ala Pro Phe Arg
                165                 170                 175

His Leu Ser Pro Ala Ser Arg Glu Gly Ala Arg Arg Leu Ala Ala Glu
            180                 185                 190

Ala Glu Leu Ala Leu Ser Gly Arg Thr Trp Ala Pro Gly Val Glu Ala
        195                 200                 205

Leu Thr His Thr Leu Leu Ser Thr Ala Val Asn Asn Met Met Leu Arg
    210                 215                 220

Asp Arg Trp Ser Leu Val Ala Glu Arg Arg Arg Gln Ala Gly Ile Ala
225                 230                 235                 240

Gly His Thr Tyr Leu Gln Ala Ser Glu Lys Phe Lys Met Trp Gly Ala
                245                 250                 255

Glu Pro Val Ser Ala Pro Ala Arg Gly Tyr Lys Asn Gly Ala Pro Glu
            260                 265                 270

Ser Thr Asp Ile Pro Pro Gly Ser Ile Ala Ala Ala Pro Gln Gly Asp
        275                 280                 285

Arg Cys Pro Ile Val Arg Gln Arg Gly Val Ala Leu Ser Pro Val Leu
    290                 295                 300

Pro Pro Met Asn Pro Val Pro Thr Ser Gly Thr Pro Ala Pro Ala Pro
305                 310                 315                 320

Pro Gly Asp Gly Ser Tyr Leu Trp Ile Pro Ala Ser His Tyr Asn Gln
                325                 330                 335

Leu Val Ala Gly His Ala Ala Pro Gln Pro Gln Pro His Ser Ala Phe
            340                 345                 350

Gly Phe Pro Ala Ala Ala Gly Ser Val Ala Tyr Gly Pro His Gly Ala
        355                 360                 365

Gly Leu Ser Gln His Tyr Pro Pro His Val Ala His Gln Tyr Pro Gly
    370                 375                 380

Val Leu Phe Ser Gly Pro Ser Pro Leu Glu Ala Gln Ile Ala Ala Leu
385                 390                 395                 400

Val Gly Ala Ile Ala Ala Asp Arg Gln Ala Gly Gly Gln Pro Ala Ala
                405                 410                 415
```

```
Gly Asp Pro Gly Val Arg Gly Ser Gly Lys Arg Arg Arg Tyr Glu Ala
          420                 425                 430

Gly Pro Ser Glu Ser Tyr Cys Asp Gln Asp Glu Pro Asp Ala Asp Tyr
      435                 440                 445

Pro Tyr Tyr Pro Gly Glu Ala Arg Gly Ala Pro Arg Gly Val Asp Ser
  450                 455                 460

Arg Arg Ala Ala Arg His Ser Pro Gly Thr Asn Glu Thr Ile Thr Ala
465                 470                 475                 480

Leu Met Gly Ala Val Thr Ser Leu Gln Gln Glu Leu Ala His Met Arg
                485                 490                 495

Ala Arg Thr Ser Ala Pro Tyr Gly Met Tyr Thr Pro Val Ala His Tyr
            500                 505                 510

Arg Pro Gln Val Gly Glu Pro Glu Pro Thr Thr Thr His Pro Ala Leu
      515                 520                 525

Cys Pro Pro Glu Ala Val Tyr Arg Pro Pro Pro His Ser Ala Pro Tyr
  530                 535                 540

Gly Pro Pro Gln Gly Pro Ala Ser His Ala Pro Thr Pro Pro Tyr Ala
545                 550                 555                 560

Pro Ala Ala Cys Pro Pro Gly Pro Pro Pro Pro Pro Cys Pro Ser Thr
                565                 570                 575

Gln Thr Arg Ala Pro Leu Pro Thr Glu Pro Ala Phe Pro Pro Ala Ala
            580                 585                 590

Thr Gly Ser Gln Pro Glu Ala Ser Asn Ala Glu Ala Gly Ala Leu Val
      595                 600                 605

Asn Ala Ser Ser Ala Ala His Val Asp Val Asp Thr Ala Arg Ala Ala
  610                 615                 620

Asp Leu Phe Val Ser Gln Met Met Gly Ala Arg
625                 630                 635
```

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CCGGTGCCCA ATCGTCCGT 19

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GTCCGTGCGC GTCAAGTGG 19

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TTCCGGCTCC CCCACCTGA 19

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

ATTCGGATCC TGGAGGCGA 19

CLAIMS

1. An essentially pure protein encoded by an HSV-2 UL26 gene and functional fragments thereof.

2. The essentially pure protein of claim 1 wherein said protein is selected from the group consisting of HSV-2 protease precursor protein, mature HSV-2 protease and functional fragments of said mature HSV-2 protease.

3. The essentially pure protein of claim 1 wherein said protein is mature HSV-2 protease.

4. An essentially pure protein encoded by HSV-2 UL26.5 gene or fragments thereof.

5. The essentially pure protein of claim 4 wherein said protein is selected from the group consisting of HSV-2 capsid precursor protein, mature HSV-2 capsid protein and functional fragments thereof.

6. The essentially pure protein of claim 1 wherein said protein is mature HSV-2 capsid protein.

7. An isolated nucleic acid molecules comprising an HSV-2 UL26 gene or functional fragments thereof.

8. The isolated nucleic acid molecule of claim 7 comprising a nucleotide sequence of SEQ ID NO:1 or a functional fragment thereof.

9. The isolated nucleic acid molecule of claim 7 comprising a nucleotide sequence that encodes mature HSV-2 protease.

10. The isolated nucleic acid molecule of claim 7 comprising an HSV-2 UL26.5 gene or a functional fragment thereof.

11. The isolated nucleic acid molecule of claim 10 comprising a nucleotide sequence that encodes mature HSV-2 capsid protein.

12. The isolated nucleic acid molecule of claim 10 comprising the HSV-2 UL26.5 promoter.

13. An expression vector comprising an HSV-2 UL26 gene or functional fragment thereof.

14. The expression vector of claim 13 wherein said UL26 gene is disclosed in SEQ ID NO:1.

15. The expression vector of claim 13 wherein said fragment of said UL26 gene is selected from the group consisting of: a nucleotide sequence that encodes mature HSV-2 protease, a nucleotide sequence that encodes mature HSV-2 capsid protein, a nucleotide sequence that encodes an HSV-2 UL26.5 gene, a

nucleotide sequence that encodes mature HSV-2 capsid protein and the HSV-2 UL26.5 promoter.

16. A host cell that has been transformed with an expression vector of claim 13, said host cell being capable of expressing said UL26 gene or functional fragment thereof.

17. A method of identifying compounds that inhibit HSV-2 protease activity comprising the steps of:

a) contacting HSV-2 protease or functional fragment thereof with an HSV-2 protease substrate in the presence of a test compound;

b) detecting the level of proteolytic cleavage of said substrate; and

c) comparing that level to the level of proteolytic activity that occurs when HSV-2 protease or functional fragment thereof is contacted with an HSV-2 protease substrate in the absence of a test compound.

18. A method of identifying compounds that inhibit HSV-2 virion assembly comprising

a) in the presence of a test compound, contacting two or more proteins that comprise at least portions of HSV-2 capsid protein in the presence of a test compound;

b) detecting the level of capsid-capsid association; and

c) comparing said level of capsid-capsid association to the level of capsid-capsid association that occurs when two or more proteins that comprise at least portions of HSV-2 capsid protein are contacted in the absence of the test compound.

19. A synthetic HSV-2 protease substrate having the formula R_1 - SEQ ID NO:3 - R_2 or R_1 - SEQ ID NO:4 - R_2.

20. The synthetic HSV-2 protease substrate of claim 19 selected from the group consisting of: SEQ ID NO:3; SEQ ID NO:4; SEQ ID NO:5; SEQ ID NO:6; SEQ ID NO:7; SEQ ID NO:8; SEQ ID NO:9; SEQ ID NO:10; SEQ ID NO:11; SEQ ID NO:12; SEQ ID NO:13; SEQ ID NO:14; and SEQ ID NO:15.

21. An antibody that selectively binds to an unprocessed HSV-2 protease wherein said antibody is incapable of binding to a processed HSV-2 substrate.

22. A method of distinguishing between HSV-1 DNA and HSV-2 DNA comprising the steps of:

a) amplifying DNA in a sample using primers which amplify HSV-1 DNA but which do not amplify HSV-2 DNA and/or amplifying DNA in a

sample using primers which amplify HSV-2 DNA but which do not amplify HSV-1 DNA;

b) detecting the presence of amplified DNA.

23. A set of PCR primers comprising nucleotide sequences which can be used to amplify HSV-1 DNA but cannot be used to amplify HSV-2 DNA or comprising nucleotide sequences which can be used to amplify HSV-2 DNA but cannot be used to amplify HSV-1 DNA.

24. A kit for distinguishing between HSV-1 DNA and HSV-2 DNA comprising a container comprising a set of PCR primers of claim 23 and a container comprising a DNA size marker molecule.

25. A method of distinguishing between HSV-1 protein and HSV-2 protein comprising the steps of:

a) performing an immunoassay using antibodies capable of selectively binding to HSV-1 protein and incapable of binding to HSV-2 protein and/or performing an immunoassay using antibodies capable of selectively binding to HSV-2 protein incapable of binding to HSV-1 protein; and

b) detecting the presence of bound antibodies.

26. An antibody capable of selectively binding to HSV-2 protein and incapable of binding to HSV-1 protein and an antibody capable of selectively binding to HSV-1 protein and incapable of binding to HSV-2 protein.

27. A kit for distinguishing between HSV-1 protein and HSV-2 protein comprising a container comprising an antibody of claim 26 and/or a container comprising an antibody capable of selectively binding to HSV-1 protein and incapable of binding to HSV-2 protein and/or antibody capable of selectively binding to HSV-2 protein and incapable of binding to HSV-1 protein.

28. HSV-2 protease inhibitor compounds identified by the method of claim 17.

29. HSV-2 virion assembly inhibiting compounds identified by the method of claim 18.

1 GTCGACGAGGCGCGTGGTGGATATGTCGTCGGGCGCCCGCCAGGCGGCGCTCGTGCGCCTCACCGCGCTGGAGCTCATCAACCGCACCCGCACAAACACC 100
101 ACCCCTGTGGGGGAGATTATTAACGCCCACGATGCCTTGGGGATACAATACGAACAGGGCCTGGGGCTGCTCGCCCAGCAGGCACGCATCGGCTTGGCGT 200
201 CGAACGCCAAGCGATTCGCCACGTTCAACGTGGGCAGCGACTACGACCTGTTGTACTTTTTGTGTCTCGGGTTCATTCCCCAGTACCTGTCCGTGGCCTA 300
301 GGGAAGGGTGGGGGTGGTGGTGGGGTGTTTTTCTGTTGTTGTTCTGGTCCGCCTGGTCACAAAGGCACGGCGCCCCGAAACGCGGGCTTTAG 400
401 TCCCGGCCCCGGACGTCGGCGGACACACAACGGCGGGCCCCGTGGGTGGGTAAGTTGGTTCGGGGGCATCGCTGTATTCCCTTGCCCGCTTCCACCCC 500
501 CCCTTCCCGTTTGGTTTGTTGTGGCGGTGCCCATGGCGTCGGCGGAAATGCGCGAGCGGTTGGAGGCGCCTCTGCCCGACCGGGCGGTGCCCATCTACG 600
M A S A E M R R E R L E A P L P D R A V P I Y V
601 TGGCCGGGTTTTTGGCCCTGTACGACAGCGGGACCCGGCGAGCTGGCCCTGGACCCAGACACGGTGCGTGCGGCCCTGCCTCCGGAGAACCCCCTGCC 700
A G F L A L Y D S G D P G E L A L D P D T V R A A L P P E N P L P
701 GATCAACGTAGACCACCGCCGCTCGGTGCGAGGTGGGCCGGGTGCTCGCCGTGGTCAACGACCCTCGGGGGCCGTTTTTGTGGGGCTGATCGCGTGCGTG 800
I N V D H R A R C E V G R V L A V V N D P R G P F F V G L I A C V
801 CAGCTGGAGCGCGTCCTCGAGACGGCCGCCAGCGCCGCTATTTTTGAGCGCCCGCGCTCTCCCGGGAGGAGCGTCTGCTGTACCTGATCACCA 900
Q L E R V L E T A A S A A I F E R R G P A L S R E E R L L Y L I T N
901 ACTACCTGCCATCGGTCTCGCTGTCCACAAAACGCCGGGGGGACGAGGTTCCGCCCGACCGCACCCTGTTTGCGCACGTGGCCCTGTGCGCCATCGGGCG 1000
Y L P S V S L S T K R R G D E V P P D R T L F A H V A L C A I G R
1001 GCGCCTTGGAACCATCGTCACCTACGACACCAGCCTAGACGCGGCCATCGCTCCGTTTCGCCACCTGGACCCGGCGACGCGCGAGGGGGTGCGACGCGAG 1100
R L G T I V T Y D T S L D A A I A P F R H L D P A T R E G V R R E
1101 GCCGCCGAGGCCGAGCTCGCCGCTGGCCGGGCGCACCTGGGCCCCCGGCGTGGAGGCGCTCACACACACGGCTGCTCTCCACCGCCGTCAACAACATGATGC 1200
A A E A E L A L A G R T W A P G G V E A L T H T L L S T A V N N N M M L
1201 TGCGTGACCGCCTGGAGCCTCGTGGCCGAGCGGCGGCAGGCCGGGATCGCCGGACACACGTAGCCTTCAGGCGAGCGAAAAATTTAAAATATGGGGGGC 1300
R D R W S L V A E R R R Q A G G I A G H T Y L Q A S E K F K I W G A
1301 GGAGTCTGCCCCTGCGCCGGAGCGTGGGTATAAACCGGCGCCCCCGGGTGCCATGGACACATCCCCCGCCGCGAGCGTTCCCGCGCCGCAGGTCGCCGTC 1400
E S A P A P E R G Y K T G A P G A M D T S P A A S V P A P Q V A V
1401 CGTGCGCGTCAAGTCGCGTCGTCGTCGTTCTTCTTTTCCGGCACCGGCCGATATGAACCCCGTTTCGGCATCGGCGCCCCGGCCCCTCCGC 1500
R A R Q V A S S S S S S S F P A P A D M N P V S A S G A P A P P P

FIGURE 1(a)

```
1501  CGCCCGGCGACGGGAGTTATTGTGGATCCCCGCCTCTCATTACAATCAGCTCGTCACCGGGCAATCCGCGCCCCGCCACCCGCCGCTGACCGCGTGCGG   1600
        P  G  D  G  S  Y  L  W  I  P  A  S  H  Y  N  Q  L  V  T  G  Q  S  A  P  R  H  P  P  L  T  A  C  G

1601  CCTGCCCGGCCGCGGGACGGTGGCCTACGGACACCCCGGCGCCCGGTCCCCGCACTACCCGCCTCCTCCCGCCCACCCGTACCCGGGTATGCTGTTC   1700
        L  P  A  A  G  T  V  A  Y  G  H  P  G  A  G  P  S  P  H  Y  P  P  P  A  H  P  Y  P  G  M  L  F

1701  GCGGGCCCCAGTCCCCTGGAGGCCCAGATCGCCGCGCTGGTGGGGGCCATCGCCGCCGACCGCGGGTGGCTTCCGGCGGCCGCCGGAGACCACG   1800
        A  G  P  S  P  L  E  A  Q  I  A  A  L  V  G  A  I  A  A  D  R  Q  A  G  G  L  P  A  A  A  G  D  H  G

1801  GGATCCGGGGGTCGGCGAAGCGCCGCCGACACGAGGTGGAGCAGCCGGAGTACGACTGCGGCCGTGACGAGCCGGACCGGGACTTCCCGTATTACCCGGG   1900
          I  R  G  S  A  K  R  R  R  H  E  V  E  Q  P  E  Y  D  C  G  R  D  E  P  D  R  D  F  P  Y  Y  P  G

1901  CGAGGCCCGCCCCGAGCCCGGCCCGGTCGACTCCCGGCGCGCCGCCAGGCTTCCGGGCCCCACGAAACCATCACGGCGCTGGTGGGGGCCGGTGACG   2000
        E  A  R  P  E  P  R  P  V  D  S  R  R  A  A  R  Q  A  S  G  P  H  E  T  I  T  A  L  V  G  A  V  T

2001  TCCCTGCAGCAGGAACTGGCGCACATGCGCGCGTACCCACGGCCCCCTACGGGCCGTATCCGCCGGTGGGGCCCTACCACCACCCCACGCAGACACGG   2100
        S  L  Q  Q  E  L  A  H  M  R  A  R  T  H  A  P  Y  G  P  Y  P  P  V  G  P  Y  H  H  P  H  A  D  T  E

2101  AGACCCCCGCCCAACCACCCCGCCTACCCCGCCGAGGCCGTCTATCTGCCGCCCGCCGCACATCGCCCCCCCGGGGCCTCCTCTATCCGGGCCGGTCCCCCC   2200
          T  P  A  Q  P  P  R  Y  P  A  E  A  V  Y  L  P  P  P  H  I  A  P  P  G  P  P  L  S  G  A  V  P  P

2201  ACCCTCGTATCCCCCAGTTGCGGTTACCCCCGGTCCCGCTCCCCCGCTACATCAGCCCTCCCCCCGCACACGCCCACCCCCCTCCGCCGCCGGGACCC   2300
        P  S  Y  P  P  V  A  V  T  P  G  P  A  P  P  L  H  Q  P  S  P  A  H  A  H  P  P  P  P  P  P  G  P

2301  ACGCCTCCCCCCCGCCGCGAGCTTACCCCAACCCGAGGCGCCGGCGGAGGCCGGCGCCTTAGTTAACGCCAGCAGCGCGGCCCACGTGAACGTGGACA   2400
        T  P  P  P  A  A  S  L  P  Q  P  E  A  P  G  A  E  A  G  A  L  [V  N  A  *S  ]S  A  A  H  V  N  V  D  T

2401  CGGCCCGGGCCGCCGATCTGTTTGTGTCACAGATGATGGGGTCCCGCTAACTCGCCTCCAGGATCCGAATTC  2472
        A  R  A  A  D  L  F  V  S  Q  M  M  G  S  R]]>
```

FIGURE 1(b)

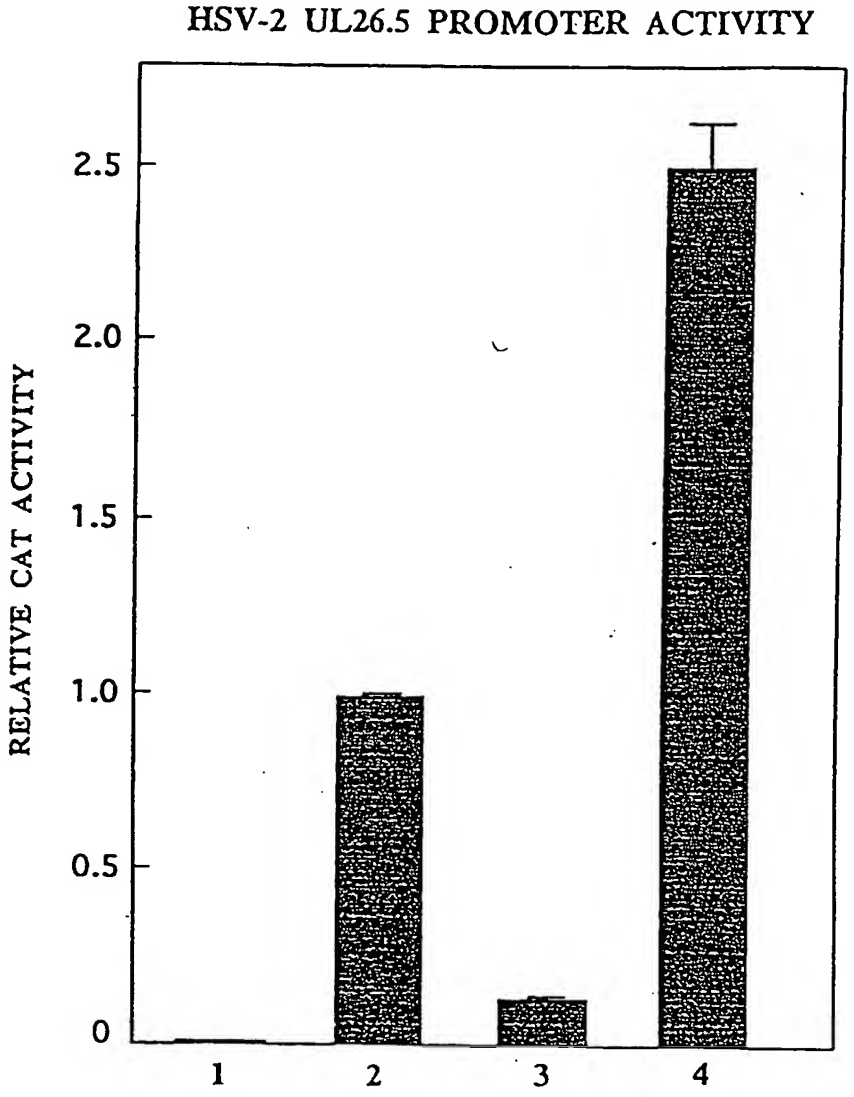
HSV-2 UL26.5 PROMOTER ACTIVITY
RELATIVE CAT ACTIVITY
2.5
2.0
1.5
1.0
0.5
0
1
2
3
4

FIGURE 2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/09303

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C07H 17/00; C12N 15/00, 5/00; C12Q 1/70; A61K 38/00, 35/14
US CL : 435, 240.1, 320.1; 530/330, 387.1; 536/23.72, 24.33; 930/220
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435, 240.1, 320.1; 530/330, 387.1; 536/23.72, 24.33; 930/220

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS; STN, FILE BIOSIS;
SEARCH TERMS: HSV-2, PROTEASE, UL26, UL26.5, CAPSID

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A, T	US, A, 5,324,664 (NUNBERG ET AL.) 28 June 1994, see entire document.	22, 23
A	US, A, 4,891,315 (WATSON ET AL.) 02 January 1990, see entire document.	1-16, 20, and 21
X	us, a, 4,709,011 (COHEN ET AL.) 24 November 1987, see columns 7 and 24.	25-27
A	US, A, 4,618,578 (BURKE ET AL.) 21 October 1986, see columns 1-6.	1-16, 20, 21
A	US, A, 5,122,449 (GILBERT ET AL.) 16 June 1992, see entire document.	1-16, 20

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search	Date of mailing of the international search report
07 NOVEMBER 1994	31 JAN 1995
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer Bradley L. Sisson Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/09303

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X, P	JOURNAL OF GENERAL VIROLOGY, Volume 74, Part 10, issued October 1993, "Herpes simplex virus type 1 capsid protein, VP21, originates within the UL26 open reading frame", pages 2269-2273, see pages 2269-2272.	1-29
X	JOURNAL OF GENERAL VIROLOGY, Volume 73, Part 10, "Identification of genes encoding two capsid proteins (VP24 and VP26) of herpes simplex virus type 1", pages 2709-2713; see pages 2709-2712.	1-29
X	JOURNAL OF VIROLOGY, Volume 65, Number 1, issued January 1991, "The Promoter, Transcriptional Unit, and Coding Sequence of Herpes Simplex Virus 1 Family 35 Proteins Are Contained within and in Frame with the UL26 Open Reading Frame", pages 206-212; see the entire document.	1-16, 20, 21, 26
X	JOURNAL OF VIROLOGY, Volume 65, Number 10, issued October 1991, "The Herpes Simplex Virus 1 Gene Encoding a Protease Also Contains within Its Coding Domain the Gene Encoding the More Abundant Substrate", pages 5149-5156; see entire document.	1-29
X	JOURNAL OF GENERAL VIROLOGY, Volume 69, issued 1988, :The Products of Herpes Simplex Virus Type 1 Gene UL26 which Are Involved in DNA Packaging Are Strongly Associated with Empty but Not with Full Capsids", pages 2879-2891; see entire document.	1-29
X, P	JOURNAL OF VIROLOGY, Volume 68, Number 4, issued April 1994, "Assembly of Herpes Simplex Virus (HSV) Intermedite Capsids in Insect Cells Infected with Recombinant Baculoviruses Expressing HSV Capsid Proteins", pages 2442-2457; see entire document.	1-29
X, P	JOURNAL OF VIROLOGY, Volume 68, Number 6, issued June 1994, "The Protease of Herpes Simplex Virus Type 1 is Essential for Functional Capsid Formation and Viral Growth", pages 3702-3712; see entire document.	1-29
T	JOURNAL OF VIROLOGY, Volume 68, Number 9, issued September 1994, "Phenotype of the Herpes Simplex Virus Type 1 Protease Substrate ICP35 Mutant Virus", pages 5384-5394; see entire document.	1-29

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/09303

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☒ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/09303

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

In the examination of international applications filed under the Patent Cooperation Treaty, PCT Rule 13.1 states that the international application shall relate to one invention only or to a group of inventions so linked as to form "a single general inventive concept." PCT Rule 13.2 indicates that this shall be construed as permitting, in particular, one of the following three possible combinations of the claimed invention:
(1) a product, a process specifically adapted for the manufacture of said product and a use of said product, or
(2) a process, and an apparatus or means specifically designed for carrying out said process, or
(3) a product, a process specifically adapted for the manufacture of said product and an apparatus or means designed for carrying out the process.
This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept. In order for all inventions to be examined, the appropriate additional examination fees must be paid.
Grouping of the Claims:
I. Claims 1-6, drawn to protein, claims 17-20, drawn to a method of identifying compounds that inhibit HSV-2 protease activity, and claims 28-29, drawn to HSV-2 protease inhibitor compounds.
II. Claims 7-12, drawn to nucleic acid.
III. Claims 13-15, drawn to an expression vector, and claim 16, drawn to transfected host cell.IV. Claim 21, drawn to an antibody that selectively binds to an unprocessed HSV-2 substrate.
V. Claims 22 and 24, drawn to a method of distinguishing between HSV-1 DNA and HSV-2 DNA and related kit, and claim 23, drawn to PCR primers.
VI. Claims 25 and 27, drawn to a method of distinguishing between HSV-1 DNA and HSV-2 protein and related kit and claim 26, drawn to an antibody capable of selectively binding HSV-2 protein and incapable of binding HSV-1 protein and an antibody capable of binding selectively HSV-1 protein and incapable of binding HSV-2 protein.

The inventions listed as Groups I-VI do not meet the requirements for Unity of invention for the following reasons:

The inventions of Group I-VI are all linked by the UL26 gene and the protein encoded thereby and expressed therefrom where said protein is a capsid protein of HSV. The publication of Liu and Roizman (JOURNAL OF VIROLOGY, January 1991) teach in detail of the UL26 open reading frame of HSV-1, the proteins encoded by sid gene and the development and use of antibodies which bind to same. Accordingly, the UL26 gene, the protein encoded thereby, and monoclonal antibodies which bind to same were all known prior to the priority date of the subject PCT application (20 August 1993) and thereby do not qualify as a special technical feature within the meaning of PCT Rule 13.2

Accordingly, the claims are not so linked by a special technicalfeature within the meaning of PCT Rule 13.2 so as to form a single inventive concept. Note that PCT Rule 13 does not provide for multiple products or methods of use within a single application.